DOCUMENT RESUME

ED 159 335 CE 016 655

AUTHOR	Quinn, Robert P.
TITLE	Effectiveness in Work Roles: Employee Responses to Work Environments. Volume I.
INSTITUTION	Michigan Univ., Ann Arbor. Survey Research Center.
SPONS AGENCY	Manpower Administration (DOL), Washington, D.C. Office of Research and Development.
PUB DATE	77
CONTRACT	DL-92-26-72-35
NOTE	946p.; For the appendixes see CE 016 610
AVAILABLE FROM	National Technical Information Service, Springfield, Virginia 22151
EDRS PRICE	MF-$1.67 HC-$50.23 Plus Postage.
DESCRIPTORS	Data Collection; Employee Attitudes; Employer Attitudes; *Evaluation Methods; Job Analysis; *Job Satisfaction; Motivation; Observation; Occupational Surveys; *Performance Factors; Personnel Evaluation; Questionnaires; *Task Performance; Vocational Adjustment; *Work Attitudes; *Work Environment
IDENTIFIERS	*Work Roles

ABSTRACT

A five-year study of effectiveness in work roles had four general objectives: (1) to assess associations between aspects of working conditions and indicators of employees' work role effectiveness; (2) to identify personal and situational characteristics that limit associations between working conditions and effectiveness; (3) to begin to map the statistical structure of associations among various classes of effectiveness measures; and (4) to assess the validity of effectiveness indicators when measured in different ways and tested against different causal factors. Major methodological findings were that information from different sources regarding working conditions and worker behavior is in agreement only for relatively unambiguous and external aspects of work. As the abstractness and, thus, potential ambiguity of a measure increase, assessments from workers, observers, and supervisors become increasingly liable to judgmental biases, such as the halo effect. Major substantive results include identification of stress effects upon worker attitudes and behaviors, including a withdrawal syndrome that begins with frequent absences and culminates in voluntary turnover. (Twenty-one chapters, each a self-contained paper including methodological information and citation of sources, comprise this final report. It is divided into two parts. Part I presents chapters on methodological topics including defining, measuring, and assessing the quality of employment, and assessing work environments with observational methods. Part II, employee responses to work environments, has three sections: work role stress and strain; motivation and rewards; and compatibility of work roles and life roles. A methodological appendix is available separately as CE 016 610.) (Author/JH)

ED159335

Effectiveness in Work Roles: Employee Responses to Work Environments

Volume I

CE 016 635

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Survey Research Center
The University of Michigan

EFFECTIVENESS IN WORK ROLES:
EMPLOYEE RESPONSES TO WORK ENVIRONMENTS

Volume I

SURVEY RESEARCH CENTER
INSTITUTE FOR SOCIAL RESEARCH
THE UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

Robert P. Quinn
Principal Investigator

This report was prepared for the Manpower Administration, U. S. Department of Labor, under research and development contract No. 92-26-72-35. Since contractors conducting research and development projects under Government sponsorship are encouraged to express their own judgement freely, this report does not necessarily represent the official opinion or policy of the Department of Labor. The contractor is solely responsible for the contents of this report.

1977

BIBLIOGRAPHIC DATA SHEET

1. Report No. DLMA 92-26-72-35-3

2.

3. Recipient's Accession No.

4. Title and Subtitle

EFFECTIVENESS IN WORK ROLES: EMPLOYEE RESPONSES TO WORK ENVIRONMENTS

5. Report Date December 31, 1977

6.

7. Author(s) Survey Research Center (Principal Investigator: Robert P. Quinn)

8. Performing Organization Rept. No.

9. Performing Organization Name and Address

Survey Research Center
Institute for Social Research
Box 1248
Ann Arbor, MI 48106

10. Project/Task/Work Unit No.

11. Contract/Grant No. DL 92-26-72-35

12. Sponsoring Organization Name and Address

U.S. Department of Labor
Manpower Administration
Office of Research and Development
1111 20th St., N.W., Washington, D. C. 20210

13. Type of Report & Period Covered Final report June 15,1972-Dec.31,1977

14.

15. Supplementary Notes

16. Abstracts This study had four general objectives: (1) to assess associations between aspects of working conditions and indicators of employees' work role effectiveness; (2) to identify personal and situational characteristics that limit associations between working conditions and effectiveness; (3) to begin to map the statistical structure of associations among various classes of effectiveness measures, and (4) to assess the validity of effectiveness indicators when measured in different ways and tested against different causal factors. Major methodological findings were that information from different sources regarding working conditions and worker behavior is in agreement only for relatively unambiguous and external aspects of work. As the abstractness and, thus, potential ambiguity of a measure increase, assessments from workers, observers and supervisors become increasingly liable to judgmental biases, such as the halo effect. Major substantive results include identification of stress effects upon worker attitudes and behaviors, including a withdrawal syndrome that begins with frequent absences and culminates in voluntary turnover.

17. Key Words and Document Analysis. 17a. Descriptors

Absenteeism
Adjustment
Attitude survey
Attitudes
Earnings
Education
Effectiveness
Employment
Evaluation
Fringe benefits
Incentives
Job Analysis
Job description
Job satisfaction
Labor
Mobility
Motivation
Occupational diseases
Performance
Performance evaluation
Productivity
Supervision
Supervisors
Surveys

17b. Identifiers/Open-Ended Terms

Work Attitudes
Work roles
Employee responses
Longitudinal surveys
Lifestyles
Work environments
Leisure activities
Family roles
Social indicators
Meaning of work

17c. COSATI Field/Group 5I, 5J, 6J

18. Availability Statement Distribution is unlimited. Available from National Technical Information Service, Springfield, Va. 22151.

19. Security Class (This Report) UNCLASSIFIED

20. Security Class (This Page) UNCLASSIFIED

21. No. of Pages 1049

22. Price

THIS FORM MAY BE REPRODUCED

PREFACE

This final report brings to a close a five year (1972-1977) research project titled "Effectiveness in Work Roles." The primary focus of this study was the impact of work environments upon workers, as viewed from the diverse perspectives of the workers themselves, the employers and the society at large. The aims, equally, were to advance the technology for the assessment of work environments, and to add substantive knowledge regarding certain consequences of work-life. The study plan included two separate phases of data collection, in 1972-73 and in 1974. The present report, which summarizes all components and phases of the study, is in two volumes.

Volume I begins with an overview of the approach used and executive summary of major findings. The chapters, twenty-one in number, spell out the details of methods used, results obtained and conclusions reached. Volume I is in two parts, the first dealing with methodological issues, and the second with substantive topics. Part I of this volume includes, first, a set of chapters attempting to validate a broad conception of the meaning of quality of employment, followed by a section on the reliability and validity of observational methods, and finally, a section on certain issues of strategy in the analysis and interpretation of data concerning quality of employment and effectiveness in work roles. Part II of Volume I also has three sections: one on work-related stress and strain, a second on work motivations and rewards, and a third on work roles in relation to the worker's other life roles.

Volume II contains nine methodological appendices that provide technical details of the measures used.

This report endeavors to be comprehensive at the cost of some redundancy. By design, each chapter is a self-contained paper, with all necessary methodological information and citation of sources. The various chapters thus, to some extent, repeat the description of methodological details. Also, some of the earlier reports by project staff are included in their original form even though their methodological sections were, in some cases, superceded by later work.

Some chapters have appeared in other sources and forms, as follows: Chapters 3 and 4 are taken from an interim report to the sponsor, dated 1975. Chapter 4 also appears in *Journal of Applied Psychology*, 1975, *60*, 171-181. Chapter 13 appears in the same journal, 1976, *61*, 35-40. Chapters 5, 6, 7 and 9 appear as part of a technical report to the National Science Foundation, 1975. Chapter 14 is a paper read before the Annual Meeting of the American Psychological Association, Chicago, 1975. Chapters 18 and 20 are abstracted from the doctoral dissertations of the respective authors.

ACKNOWLEDGEMENTS

The scope and complexity of this study required the cooperation and support of several institutions as well as a large cast of individuals. We are particularly indebted to the Manpower Administration of the Department of Labor which funded both phases of the project (under research and development contract No. 92-26-72-35). Special thanks go to Dr. Robert Foster who, as our contract officer at the Department of Labor, displayed great ingenuity and forebearance in the face of unexpected problems and delays. We also extend our appreciation to the National Science Foundation which provided supplementary (partial) funding for the analyses reported in Chapters 5, 6, 7 and 9 (Grant No. GI-29904 from the National Science Foundation's Division of Research Applied to National Needs). We thank, too, the representatives of the Field Office and of the Coding Section of the Institute for Social Research, who assisted in the data collection at both phases, and the members of the organizations that participated in the study who gave generously of their time and energy to make the study possible.

We also wish to acknowledge the varied contributions of individual members of the study staff. Robert Quinn was a principal investigator from the inception of the project; he contributed much to the original conceptual approach and guided the strategic decisions regarding evaluation of measures. Cortlandt Cammann joined the study in its early stages and heroically piloted the field operations at five sites through a stormy passage. In addition, he played a major role in several of the initial reports. Nina Gupta and Terry Beehr contributed generously to all components of the study from the beginning to end, and their impact is made evident by the large number of chapters that bear their names. Douglas Jenkins and David Nadler led the successful effort to test a measurement technology based on standardized observations, clearly one of the more important outcomes of the study.

Beyond these major investigators are the many authors of chapters in this volume whose contributions we can acknowledge only in the Table of Contents. We should also give special mention to the tireless efforts of Margaret Black and David Pagnucco who willingly assumed major responsibility for day-to-day project operations including quality control and vast quantities of data processing.

Bernard Goitein's assistance in the later stages was invaluable in helping us edit this volume. Marion Mittendorf and Carolyn Weeks efficiently organized the typing and proofing of the final manuscript.

In addition to their roles as editors of this final report, Stanley Seashore was an overseer of the project throughout most of its duration, and Graham Staines coordinated the final stages of analysis and report preparation.

TABLE OF CONTENTS

Preface iii

Acknowledgements iv

Overview and Executive Summary viii

Volume I, Part I: Methodological Topics

Section A Defining, Measuring and Assessing the Quality of Employment

Chapter 1. Data Acquisition and Field Methods - Phase I . . . 1

Chapter 2. Data Acquisition and Field Methods - Phase II . . . 19

Chapter 3. Measuring Effectiveness in Work Roles from Three Perspectives
Cortlandt Cammann, Robert P. Quinn, Terry A. Beehr, & Nina Gupta 31

Section B Assessing Work Environments with Observational Methods

Chapter 4. Standardized Observations: An Approach to Measuring the Nature of Jobs
G. Douglas Jenkins Jr., David A. Nadler, Edward E. Lawler III, & Cortlandt Cammann 141

Chapter 5. Bias Inherent in Observational Measures
Michael Moch, Cortlandt Cammann, & Nina Gupta . . . 171

Chapter 6. Obtrusive and Less Obtrustive Measures of Status, Supervisory Style, Involvement, and Job Satisfaction
Terry A. Beehr 195

Chapter 7. Standardized Observations of Job Characteristics: A Refinement and Replication
G. Douglas Jenkins Jr., & David Nadler 219

Chapter 8. Identifying Sources of Instability in Measures of Working Conditions and Work-related Attitudes and Behaviors
Bernard J. Goitein 251

Volume I, Part I: Methodological Topics (Cont.)

Section B Assessing Work Environments with Observational Methods (Cont.)

Chapter 9. Aggregating Data in Organizational Research
Nina Gupta & Robert P. Quinn 367

Chapter 10. Job Satisfaction Measurement: Normative and Individualistic Models
Nina Gupta 385

Chapter 11. The Impact of Cognitive Sophistication on Responses to an Occupational Survey
Robert P. Quinn, Graham L. Staines, Bernard J. Goitein, & David J. Pagnucco 413

Volume I, Part II: Employee Responses to Work Environments

Section A Work Role Stress, Strain and Withdrawal

Chapter 12. Role Ambiguity and Role Strain: A Search for Intervening Variables
Terry A. Beehr 459

Chapter 13. Perceived Situational Moderators of the Relationship Between Subjective Role Ambiguity and Role Strain
Terry A. Beehr 473

Chapter 14. What Makes Jobs Monotonous and Boring
Robert P. Quinn 491

Chapter 15. Predicting Withdrawal from Work: Absenteeism and Turnover
Nina Gupta & Terry A. Beehr 507

Chapter 16. Effectiveness of Withdrawal Strategies in Coping with Work Role Stress
Terry A. Beehr 549

Chapter 17. Turnover: An Inquiry Into the Effects of Job Satisfaction, Working Conditions and Quality of Job Performance
Terry A. Beehr 581

Volume I, Part II: Employee Responses to Work Environments (Cont.)

Section B Motivation and Rewards

Chapter 18. Supervisor-subordinate Similarity: A Determinant of Subordinate Ratings and Rewards
Veronica Feria Nieva 603

Chapter 19. Alternative Motivational Strategies: An Assessment of Costs and Benefits
John R. Klesh & Cortlandt Cammann 625

Chapter 20. Impact of Performance-contingent Rewards on Job Satisfaction
Nina Gupta 665

Section C Compatibility of Work Roles and Life Roles

Chapter 21. Work and Nonwork: Part I - A Review of the Literature
Graham L. Staines 693

Work and Nonwork: Part II - An Empirical Study
Graham L. Staines & David Pagnucco 729

OVERVIEW AND EXECUTIVE SUMMARY

The original prospectus for this study of effectiveness in work roles laid out four general objectives: (1) To assess some associations between aspects of working conditions and various indicators of the employee's work role effectiveness; (2) To identify some of the personal and situational characteristics that limit associations between working conditions and effectiveness; (3) To begin to map the statistical structure of associations among various classes of effectiveness measures, and (4) To assess the validity of effectiveness indicators when measured in different ways and tested against different causal factors. Methodological interests were combined with substantive interests. We aimed not only to understand better the many ways in which a worker may become effective or ineffective, but also to find improved ways to measure the indicators of effectiveness.

This overview describes how we approached the task, our perspective on what is to be meant by "effectiveness", the strategies employed in this research, and some of the main conclusions reached.

The central theme of our approach is expressed in the term "multiples". We aimed to examine the work roles of people in a variety of ordinary organizations and jobs. We aimed to measure as many diverse indicators of effectiveness as possible, and to employ multiple methods of measurement where feasible. Since the idea of effectiveness is evaluative, implying the legitimacy of different value perspectives, we attempted to include measures presumed to be of special or unique concern from these multiple perspectives. Since we view work role effectiveness as a dynamic, changing pattern of events, conditions and behaviors, we chose to obtain measures at two different times

A person's work role includes all behavioral and attitudinal aspects of the worker's life that, in any significant degree, influence or are influenced by the job or by the conditions of employment. A work role, in this view, is not abandoned daily at the factory gate or office door but is carried into other activity spheres, in other places, at other times. The role includes working, but much more. To some degree, a person's work role is potentially being enacted twenty-four hours each day, every day, and over the whole of the adult life span.

Role. The notion of "role" that we employ is a standard one: the role consists of all the behavioral requirements and expectations that are directed toward the person by significant others. In the case of a work role, these significant others obviously include supervisors, coworkers and others at the workplace itself. However, there are for most people, as well, significant role senders (and role effectiveness evaluators) away from work. Examples are the worker's family, friends and neighbors, all of whom may, by their requirements and expectations of the person, help to define that person's work role. There is also the bank officer, whose response to an application for a mortgage loan is influenced by the applicant's occupation and he, in turn, may influence the manner in which the worker fulfills that work role. People distant in time and place may become a part of the work role defining and work role evaluating set of significant others. Examples: the tax collector, the welfare officer, the person's grandchildren. The network of role senders and role performance evaluators is, in principle, never fully defined; our active interest, of course, is limited to those who are close to the scene of work role enactment and who are importantly part of the action.

with an interval of about 20 months and obtained some measures on a continuous basis. Finally, in our analytic strategies, we preferred to work not with single measures or pairs of measures, but with sets of multiple measures thought to be interdependent in some way.

To collect the data, we first interviewed at length about 650 full-time employed adults, in a wide range of common occupations, in five different employing establishments. The employers' records provided initial and continuing data with respect to such matters as pay, absences, quits and transfers, and the like. For some people and their jobs, trained observers visited the workplaces to get an independent report and description of the work processes and conditions. About 270 of the original population were included in a follow-up measurement phase, 18 to 22 months after the original measurements, when most but not all of the original procedures were repeated. Details about the population of persons, organizations and jobs, as well as the measurement methods, are provided in Chapters 1 and 2. The appendix volume contains all of the interview, questionnaire, rating, and observation documents we employed.

The Meaning of Effectiveness

The title of this study was chosen, with intent, to be somewhat ambiguous. The term "work role" is used to suggest a broader concern than is implied by such familiar words as: job, work, occupation, or employment. We use the term "effectiveness" to suggest a range of evaluation that goes beyond the direct output from job performance. The following paragraphs explain what we do mean by the phrase "effectiveness in work roles".

Effectiveness. The term "effectiveness" is intended, similarly, to be inclusive. We mean to express something more than on-job work performance, although that is prominently included. The intent is to invite consideration of a broad range of outcomes from work role enactment and a broad range of interests and value perspectives from which these outcomes can be evaluated.

The effectivensss of a person's work role performance inevitably will be judged from different value perspectives. The evaluators can be roughly classified into three broad interest-group categories: (1) the employer, (2) the general public or society, and (3) the worker himself along with others who share his personal concerns.

The employee's own judgment of work role effectivenss is made with referrence to all of the personal needs, aspirations, fears and pleasures thought to be in any way related to work role outcomes. The employer or employing organization will invoke a unique set of value priorities and specific, desired work performance outcomes that are to be satisfied through the worker's role performance. The public or societal interests arise from the needs of persons, organizations and social institutions that may be peripheral to the worker's work role but that are in some way interdependent with the person's work role enactment and its consequences; that is, we recognize that some work role outcomes (e.g. injury, layoff, promotion, efficient production) are threatening or supportive to the general social well being, and are to be evaluated in such terms.

It is important to note that these three general categories of interest perspective are not themselves entirely consistent or coherent. The employer may need high volume productivity today but improved quality tomorrow; the immediate supervisor may need to retain a highly productive worker, but the organization may need his transfer or promotion. Similarly, the wife and husband may share generally in their concerns about the worker's work role effectiveness, but may differ in their priorities as between, say, more pay or better working hours.

Time. We now mention again, for emphasis, the factor of time in work role effectiveness. The enactment of a person's work role must be evaluated not only with reference to immediate or short-run outcomes, but as well with reference to delayed consequences -- some of which may extend, life-long, beyond the period of active employment into retirement and into the lives of survivors. Consider the example of a coal miner whose present black lung disease shortens a future productive life, or the person whose work role effectiveness includes the successful provision of education or business capital to his children.

Structure. Finally, this explanation of our meaning of the title of report must introduce the idea of structure in role effectiveness. If the diverse value perspectives of many role evaluators are invoked, and if both immediate and longer-term effects are to be considered, then a roster of role effectiveness indicators would be very long and never complete. The conception of effectiveness would be made useless by its own complexity -- useless as a guide to thought and useless as a guide to measurement operations. This consequence, however, need not follow from the conceptual definitions. We assert that there exists a "structure" of work role effectiveness that transforms an impossibly complex array of indicators into an ordered domain of information that is susceptible to analysis and description in relatively simple terms.

By structure we mean only that we expect the multitude of potential effectiveness indicators to be not a random roster, but to be clustered, hierarchically organized, and causally linked. By clustered it is meant that certain work role enactment outcomes tend to occur together and to comprise a set of compatible, related outcomes; such clusters can be named, measured and described in compact and simplified ways. By hierarchically organized, it is meant that some role outcomes are of an inclusive sort, comprising numerous lesser outcomes subsumed under the more general rubric. By causally-linked, it meant that some "near" outcomes are instrumental -- i.e., causes, necessary conditions, or means -- for the realization of more "distant" outcomes in such a way that their orderly relationships may be mapped as causal chains or nets. These three asserted properties of the domain of work role outcomes imply that the enlargement of the scope of study to include a broad range of effectiveness indicators does not preclude efficient empirical study. Also, these assertions define some of the specific methodological and substantive research tasks reported in this volume.

Choice of Effectiveness Measures

The foregoing essay on the meaning of work role effectiveness sets the framework for understanding the choices made, in this study, of indicators of effectiveness. All together, about 200 effectiveness measures were obtained. Some of these indicators are used in their original form, while others were treated as clusters subsumed under a more general label. An example is the measure of annual rate of pay, which in some instances is used as a separate measure of a work role outcome valued by the worker, but occasionally is merged with other indicators to form a more inclusive indicator of "economic benefits from work".

The roster of effectiveness measures was chosen to be diverse, within limits of convenience, cost and feasibility. For example, we employ some criteria of effectiveness that are conceptually, physically or in time separated from the workplace: depressed mood, life satisfaction, social engagement with friends and neighbors. Some criteria represent the individual worker's immediate on-the-job interests and needs: work role strain, job satisfaction, intention to quit the job. Other criteria emphasize the value and interest perspective of employers: work motivation, rated productivity, abseentism, turnover. Societal interests are represented, although rather sparsely, by such measures as workers' political participation and membership in non-work organizations. The effectiveness indicators include some that represent delayed consequences of work role behavior of kinds that can be detected over a span of two years: changes in absence rates, reduction or increase in work role strain, voluntary quits.

A given effectiveness indicator may mean different things from different value perspectives, depending upon the context of its use and the forms of analysis. An example will illustrate this point. We recorded voluntary absences as an actual record of scheduled days of work missed by each employee. From the employers' viewpoint, the interest might well be in the total number of days of absence in a given period, each day representing about the same implications as to cost, time loss, and schedule accommodation. However, from the individual employee's point of view, and for analysis of the antecedents and causes of voluntary absence, a better measures is the number of occasions of absence, counting any multiple serial days of absence as a single occasion. Similarly, the employer is likely to be concerned about all instances of turnover.

regardless of the causes, while for understanding job-related causes of turnover it is useful to distinguish between voluntary and involuntary instances because they are differently caused.

The Structure of Effectiveness

The broad array of effectiveness measures obtained allowed us to explore the structure of effectiveness and also to exploit that structure for economy of analysis.

Chapter 3 reports an effort to construct a simplified model of the whole domain of effectiveness measures. It is shown that the patterns of correlation among a select set of diverse measures have a component of universality -- a patterning that appears to be stable across subpopulations of different age, sex, and family status. At the same time, there appear to be some differences among such population categories: the domain of effectiveness is more richly interconnected for younger women than for, say, older men. Conclusion: There are identifiable stable clusters of work role effectiveness criteria, but also some independent domains of effectiveness; some degree of simplified representation is feasible, but it is not feasible to represent effectiveness in a single, inclusive evaluative measure.

Another aspect of the structure of effectiveness is the determination of sets of effectiveness measures that are homogeneous as to their causal linkages, or their intercorrelations, or both. An illustration of the conceptual and strategic issues is provided by a pair of chapters (12 and 13) dealing with work role ambiguity (a characteristic of the job) as a source of work role strain (an individual-level consequence). It is shown that the several measures of work-role strain are only modestly intercorrelated, leaving uncertain the question whether these strains are to be treated separately, as different phenomena, or whether they are alternative aspects of a single phenomenon and therefore combinable into a joint index. An exploration of possible causal linkages, however, showed that all of the strain measures appear to have about the same pattern of relationships with role ambiguity (presumed to be a partial cause of strain) and with various intervening and situational factors that might moderate the ambiguity-strain connection. Conclusion: It appears that this set of work role strain outcomes may be assessed either jointly or separately, as they display both clustering and common causal linkages.

The design of the study provided for measurements over a two year interval primarily to allow for some assessment of the stability of measurements over time, but also to allow the analytical use of changes. Three chapters use measures of deferred consequences of earlier events and conditions. Chapter 15, for example, examines the effects of initial work role stresses upon subsequent liklihood of absence and turnover; it is concluded that stressful jobs induce elevated absence rates as well as earlier and more frequent voluntary quits, and that these relationships persist even when certain contextual factors are neutralized. Chapter 16 attempts, although with limited success, to determine whether employees initially reporting work role strains in highly stressful jobs could successfully moderate the experienced strain by psychological or physical withdrawal from the work role. Withdrawal was measured in terms of absences, lateness, voluntary turnover, reduced interaction with others at work, reduced psychological involvement in the work. Although the evidence is weak, it is suggested that work role strain induces increased rates or degrees of withdrawal, but that the withdrawal does not mitigate the strain. Chapter 17 shows that relatively low-quality working conditions and job dissatisfaction at the time of initial measurement are associated with later intention to quit and liklihood of actually quitting, but that the causal mechanisms appear not to work selectively in a way to drive out the better-performing workers.

The number of effectiveness measures employed in this study precludes itemizing here all of their various uses. It should be noted that priority attention has been given to such obviously ineffective work role behaviors and outcomes as: (1) Those associated with ill health, i.e. work related strain, depresssed mood; (2) Those associated with costly events, i.e. voluntary quitting, absence, lateness, poor job performance, and (3) Those associated with loss of off-the-job social integration of the worker.

Multiple Measures

In addition to the foregoing multiplicity of effectiveness indicators and value perspectives, the design of this study gave explicit attention to assessing the validity of information about jobs, work environments, personal attributes of workers and their work-related behaviors. When feasible, measures were obtained by multiple means, and from multiple sources. The reason for this was an awareness that familiarity with accustomed measures breeds both undue contentment and undue contempt. The airline pilot knows that movements around the earth subvert the various meanings of clock time; the plumber knows that measured outdoor air temperature is, at best, a crude indication of effective heat loss potential. "Effectiveness" and "working conditions" are, no less than time and temperature, subject to some error of measurement and, more importantly, to some discrepancy between the constructs to be represented and the operations for doing so. We obtained information from four sources: individual workers, their supervisors, the employer's records, and trained observers. Seven of the chapters in this report treat the degree of convergence between measures of "the same" construct obtained from different sources; additional chapters deal with converbence of alternative measures obtained from the same source.

Multiple Sources

Jenkins and his colleagues (Chapters 4 and 7) consider workers' and observers' reports of task characteristics and find moderate convergence between the two sources of information. Jenkins and Nadler (Chapter 7) argue that "objective" task characteristics interact with the desires and goals of the individual job occupant to produce the "directly experienced" job characteristics; where they differ, both may be equally valid, but for different uses. Both the interview and observer sources of information are required for job redesign efforts in practice, and for the understanding of the relationships between workers and work environments. Beehr (Chapter 6), similarly, finds adquate convergence between observer and worker ratings of the mental skill required by the job and, between company records of worker income and the workers' comparable interview reports. In Chapter 17, Beehr finds moderate convergence between worker-reported intention to look for a new job and actual subsequent voluntary turnover. In Chapter 3, Cammann, Quinn, Beehr and Gupta present data suggesting a limited convergence, at best, between records data indicating certain jobs to be "the same" (identical supervisor and census occupation code) and workers' descriptions of working conditions. Differences between jobs as indicated by data of record were only moderately related to differences in worker reports of working conditions; substantial disagreement was found among workers presumed to be rating the "same" jobs.

Beehr (Chapter 6) finds little convergence between interview data and observer ratings with respect to more subtle factors such as supervisory style and worker affect. His result is clarified by Jenkins and his colleagues who find that observers cannot even agree among themselves in their ratings of such slightly less obvious job characteristics as job predictability, required cooperation, and dependence on others. Similarly, little agreement emerges among the observers in their assessments of such variables as required effort, worker flexibility, and adequacy of work resources; since the observers cannot agree among themselves, they cannot be expected to agree with workers' assessments of these variables from their own unique perspectives.

Aside from the convergence, or lack of it, shown by direct correlation of alternative measurement operations and sources, it is necessary also to consider the dynamic equivalence of the compared measures. That is, if two measures are to be considered the same, they should be capable of showing similar causal origins and consequences. Examples of such a test of convergence are provided in Chapter 15 and 16. Absenteeism is measured by rates derived from employers' records and also by self-report of absences in interview; these two measures of absenteeism display quite similar patterns of correlation with job and work-environment variables presumed to be causes of high absenteeism. This can be taken as strong evidence for the equivalence of the alternative measures. A contrary example appears in the case of lateness, measured by self-report and from supervisors' reports of late arrival. In this case, certain presumed causes of increased latenss (work role stresses) are shown by Beehr to have a somewhat greater impact upon the supervisors' report of subordinates' lateness than upon the subordinates' own self report. It cannot be said from information available that one measure is more verdical than the other, but they certainly are "behaving" differently in hypothetical causal relationships. Gupta and Beehr (Chapter 15) report that workers' verbal intentions to quit, and incidents of actual quitting, are not dynamically equivalent indicators of some propensity for turnover, for they are differently linked with age and job tenure. Specifically, older workers, who more often intend to quit, are not so likely as younger workers actually to do so. Plausible interpretations of this anomaly are readily provided, but it is plain that the two measures do not fully converge in their meaning.

In general, the results concerning the equivalence of information from different sources show moderate convergence with respect to salient and unambiguous characteristics of jobs, workers, and work environments. Insofar as the measures endeavor to tap more subtle features of the job, or worker responses to it, the independent sources do not converge with information supplied by the workers themselves.

It is important to note that the discrepancies between information from the different sources regarding the more salient and external aspects of the worker and the job are typically meaningful and of interest in their own right. The appendix to Chapter 3 offers a demonstration of such "meaningful divergence." Quite good overall agreement is found between worker and observer assessments of the presence or absence of particular unhealthy or hazardous working conditions. Such discrepancies as found between the observers' and workers' assessments are, however, consistently in the form of a hazard noted by an observer but not reported by the worker. Several explanations are offered to account for this discrepancy. One explanation is that workers in time become accustomed to the hazards on the job and become less likely to report such hazards or, probably, even to notice them. The observers are new to these jobs and are particularly likely to notice, for example, excessive noise at the workplace. Whether observers or workers can be regarded as the more accurate, then, is not clear, but the discrepancies between their assessments are meaningful and point to the need for both sources of information.

Within-Source Measurement Bias

The counterpart of moderate-to-low convergence, between measures intended to be equivalent in meaning, is the excess of spurious convergence between measures intended to be non-equivalent. This occurs most often when representations of different abstract constructs are sought from the same information source. A common form of such bias is the simple "halo effect", or the generalization of some prominent attribute of a person (or thing, or situation) in a way that distorts the representation of other attributes. Halo effects have been discussed recently in the psychological literature as instances of "implicit theories" that people hold about others (e.g. if a man is honest, he is probably trustworthy). The presence of halo effects and implicit theories are suggested by the authors of Chapters 3-7 and 8 as relevant to interview measures, observer ratings and supervisory ratings of subordinates. Moch, Cammann and Gupta (Chapter 4) suggest that the halo effect increases, for both observer ratings and interview responses, as the concreteness of the variable decreases. Their suggestion is consistent with the previously noted increasing convergence of data from different sources as the concreteness of the concept increases. Quinn, Staines, Goitein and Pagnucco (Chapter

11) present data which suggest, unexpectedly, that implicit theories that bias workers' self reports of working conditions are stronger among the more educated and more intelligent workers.

Nieva (Chapter 18) shows that one source of bias in supervisory ratings of subordinates is the degree of similarity of supervisor and subordinate -- similarity on a set of personality and attitudinal dimensions of no evident relevance to task performance. Nieva finds that, on the average, the greater the similarity of supervisor and subordinate, the more favorable the supervisory ratings.

Quinn et al (Chapter 11) find another kind of bias in interview data, one that affects the ratings of less educated and less intelligent workers. Such cognitively unsophisticated people are liable to confound how much they feel they receive, of a particular job facet, with how much of the facet ideally there should be in a job. The authors attribute such confusion to a preference for concrete over abstract cognitions which prevents some workers from interpreting the importance ratings as referring to hypothetical jobs, hence they respond in terms of the amount of the facet their present job concretely provides. Since the more sophisticated respondents may be expected to have a greater capacity for abstraction their ratings of how much they perceive of a particular facet in their jobs, and how important that facet is for them on any job may be expected (and, in fact, are) more independent of each other.

Multiple Measures, Multiple Sources

In summary of the foregoing pages, it can be said that this study, by intention, allowed the exploration of the equivalence of alternative ways to measure variables relating to work role effectiveness and of alternative information sources. The results are a mixed bag of reassuring convergences and some unsettling cases of bias, distortion or unexpected non-convergences that require not only improvements in the technology of measurement, but also require theoretical clarification of the meaning of divergence of measures.

Multiple Time Points

A persistent problem in studies of worklife is that of distinguishing "real" change over time from spurious change or non-change arising from measurement methods. The data for this study were obtained at two different times to allow examination of both kinds of time-dependent phenomena. In some chapters (15, 16 and 17) there are reports of the prediction of subsequent worker behavior from prior conditions. In Chapter 18, the two-time measures are employed to determine the direction of causation, which would otherwise have remained uncertain. The main analyses, however, are concerned with the equivalence of descriptive information obtained at different times.

In Chapter 8, Goitein reports high convergence (agreement, equivalence) in descriptions of working conditions obtained at different times, but considerably less convergence for attitudes and behaviors reported by workers in interview. The convergence for descriptions of working conditions, however, occurs not so much because workers remain in the same jobs as because the information is obtained from the same workers.

The evidence supporting this conclusion is available because, in the interval between the two measurements, there were a number of quits, transfers and new hires. Information about the same job could be obtained from different workers; information from the same workers could be obtained regarding the initial job and also the one transferred into. For workers transferred to new jobs, the convergence of their descriptions for the new and prior jobs was almost as strong as the convergence for those who remained in the same job for both measurements. Conversely, little convergence is found between the original descriptions and those of the workers who replaced departing or transferred workers. These data are consistent with the result mentioned earlier with respect to low convergence (same time) among workers in jobs that are defined by the employers' records to be "the same"; such "same" jobs are not perceived by their occupants to be the same. The description of jobs and working conditions (within the range of jobs included in this study) appears to be more a function of the person reporting them than of the "official" or "objective" environment. This result accentuates, but does not resolve, the question as to whether the workers' perceptions are more potent than the objective situation in the prediction and assessment of effectiveness in work roles.

High convergence of measures over a time interval implies not only equivalence of scale values of the separate measures but also convergence as to the interrelationships among the measures. This aspect of convergence is examined by Goitein with respect to 23 demographic and work-related variables for a subpopulation of workers who were in the same jobs at both times of measurement. High convergence was found, except for one variable, job tenure. This deviant variable is understandable when one considers that tenure in new jobs is typically only two to three years, and for this analysis those who were "new comers" at time 1 had become "oldtimers" by Time 2. The other variables displayed high stability of interrelationships over the two-year period. It is relevant to note here that other analyses (Chapters 4, 6 and 11) successfully explore relational stability over the time interval, even when including newcomers in the Time 2 analyses.

To find out how convergence over time in measures of working conditions and attitudes might be a function of respondent characteristics, convergence estimates were calculated for subpopulations based on age, job tenure, income, sex, educational attainment, ethnicity and gross occupational category (Chapter 8). The results are complex, but it appears that convergence of descriptions of working conditions is generally higher for people in relatively high income job and clerical jobs; also, respondents of relatively greater education display greater convergence, as do males. The results were somewhat different for descriptions of own attitudes and behavior. Conclusion: convergence over time is associated with respondent and job characteristics, but not in any easily interpretable pattern. Income, job level, and educational level are probably the key factors.

There occurs a good deal of debate, but not much empirical study, of the validity of retrospective data obtained in interviews and questionnaires. Quinn, et al. (Chapter 11) compare the respondents' Time 2 retrospective judgments of changes in their working conditions with measured changes derived from concurrent measures at Time 1 and Time 2. The validity of retrospective estimates is found to be strongly influenced by the respondents' intelligence and educational level; cognitive sophistication appears to be the key factor.

The availability of measurements at two points in time make it possible to investigate the direction and magnitude of causal relationships. Gupta and Beehr (Chapter 15), and Beehr (Chapters 16 and 17) report the effects of work-related variables upon absenteeism and turnover subsequent to the interview. They demonstrate a process of withdrawal over time, first appearing as increased absenteeism, and later as leaving the job (turnover). Nieva (Chapter 18) uses advanced techniques of analysis of causality to clarify the association between supervisor-subordinate similarities, on the one hand, and the rewards and esteem received by the subordinates. She shows that similarity induces greater rewards though the reverse process also operates.

Work and Nonwork

The final chapter of this volume deals with an aspect of our conception of the meaning of "effectiveness in work roles" that has been resistant to empirical study, namely, the extent and manner in which a person's job and work environment impact upon other life roles. Of particular interest in this chapter is a debate over two rival hypotheses: the "spillover" hypothesis and the "compensatory" hypothesis. The former argues that the worker's experiences, attitudes, skills, and styles of activity are carried over into the nonwork arena in such a way that there are similarities in the patterning of work and nonwork life. The latter argues that the work situation is likely to be deficient in need fulfillment, at least in some respects, for most workers and that they will compensate for these deficiences in their choices of leisure and family activities.

Data from the present study support the spillover hypothesis for the most part. Such support is reflected in the positive correlations between degree of involvement in work (measured subjectively) and degree of involvement in nonwork. Support is also shown in the positive correlations between general types of activities engaged in at work and a corresponding set of types of activities in nonwork. The one exception to this overall pattern of spillover concerns physical effort on the job. Workers who expend a relatively great amount of physical effort at work are less involved in their nonwork activities.

Chapter 1

DATA ACQUISITION AND FIELD METHODS - PHASE I

Chapter 1

DATA ACQUISITION AND FIELD METHODS - PHASE I

The following pages provide an overview of the methods used for population sampling and data acquisition. The account will be brief because all of Part I of this volume concerns methodological issues, and each chapter provides the necessary further details as they become relevant. These chapters will be cross-referenced in instances where they specifically detail the field procedures used. The methodological appendices in Volume II contain facsimile copies of the major printed instruments employed.

The sample for Phase I of the study was drawn from five employing organizations, from 34 different departments in these organizations, and by sampling procedures that resulted in the acquisition of usable data from 651 persons. These respondents did not, strictly speaking, constitute a sample of any defined base population although that term will be used occasionally in this report. However, the sample was chosen in ways intended to include a diversity of people in a considerable variety of common occupations within common kinds of work establishments.

The data acquisition methods included personal interviews, questionnaires, search of the employers' records, supervisors' ratings of their subordinates, and systematic on-the-job observation by trained observers.

The major part of the data acquisition for Phase I took place during the winter months of 1972-1973.

These matters are detailed in the following pages. Comparable information with respect to Phase II of the study appears in Chapter 2.

Sample

A large number of midwestern firms were invited to participate in the present study. Most declined, but five establishments did agree to participate: a hospital, a printing company, two manufacturers of automobile accessories, and a research and development company. People working less than 20 hours per week were excluded. Certain departments in each establishment, rather than the entire establishment, were used as the units from which the sample was drawn. The major criteria for selection of departments were: access granted by management, having enough respondents available for interview and observation to form a usable subsample, containing jobs with a wide range of job characteristics, and feasibility of on-the-job observations. Within each of the departments selected, all supervisory personnel were included in the sample;[1] nonsupervisory employees, who of course outnumbered supervisors, were systematically sampled at lower rates dictated by a balance between the advantages of a larger sample and cost considerations. Subsequently, those employees who felt that they had not been on their present jobs long enough to answer the job-related questions adequately were screened out by the interviewers.

---

[1] All supervisors were included in the sample because performance ratings of each respondent by his or her supervisor were needed. This was one way to insure that each respondent's supervisor was in the sample. Information regarding the supervisory status of the employee was provided by the organizations studied.

The departments selected in the hospital were those in charge of patients' nutrition, general services (custodians, elevator operators, ambulance drivers, etc.), respiratory therapy, and the laboratory services for the internal medicine department. Because the first two departments were quite large, 50 percent of nonsupervisors and 100 percent of the supervisors were included.[2] This resulted in a total of 54 supervisors and 80 nonsupervisors in the patient nutrition department and 17 supervisors and 61 nonsupervisors in the general services department.

All supervisors and nonsupervisors from the respiratory therapy and the laboratory research section of the internal medicine department were included in the sample, primarily because these departments were relatively small in number--54 employees in the first and 13 in the second department. Doctors and nurses were excluded from the sample because of the problems anticipated in attempting on-the-job observations of their behaviors. Technical personnel, respiratory therapists, and laboratory technicians were all included in the sample.

At the printing company all departments were sampled with two exceptions: the top level managers (who also owned the firm), and the sales department, which was spread across the country. Four small departments--maintenance, finance, personnel, and engineering--were sampled at a rate of 100 percent among both supervisors and nonsupervisors. All the supervisors, and 66.7 percent of the nonsupervisors were sampled[3] from the remaining departments, viz., sales services, warehouses, and eight departments dealing directly with various stages of production. Thus, the sample for the printing company included all 42 supervisors and 177 of 272 nonsupervisors in the departments selected for study.

---

[2] The sample was drawn by choosing every second employee from a list furnished by the hospital. Whether the first or second individual on the list was the first respondent chosen was determined randomly.

The sample at the first automotive supply organization was drawn from the basic production departments. The major exclusions were the sales, research and development, and accounting units. Within the sampled units, 100 percent of the supervisors and 50 percent of the nonsupervisors were included in the sample. The resulting sample included 30 supervisors and 138 non-supervisors.

At the second automotive supplier, all 15 supervisors in seven production departments were included. Among non-supervisors, 180, or 38.1 percent, of the 473 employees were included in the sample. The departments were sampled at different rates, ranging from 25 percent to 40 percent.

There were 33 employees in the research and development company. Of these, the president, vice-president, and head of the sales group were excluded from the sample because they traveled frequently; one person was excluded because of short job tenure. The remaining 29 employees were included in the sample.

Response Rates. The interview response rates were 76.2% at the hospital, 79.0% at the printer, 71.4% at the first automotive supplier, 63.6 % at the second automotive supplier, and 72.4% at the research and development company. The overall response rate for the five organizations was 72.9%. The final analysis sample consisted of 651 employees.

---

[3] The sample was systematic, e.g., 110110110, etc.

Demographic and Occupational Constitution of the Sample. The demographic and occupational constitution of the sample is reported in Table 1, together with comparative statistics from a 1973 national sample of workers. Since the sample was not intended to be representative of the national population, the differences between it and the national sample are not surprising, but it may be important to keep these differences in mind when the results of the study are interpreted. Demographically, the sample had a higher proportion of blacks, women, single people, and young people than the comparison national sample. Since the sample was drawn from five organizations performing rather specific operations, the occupational subgroups represented in it could be expected to be quite different from those in the national sample. Given this expectation, the percentages in each category for the two samples were remarkably similar. The major differences were that the Effectiveness in Work Roles' sample included a larger percentage of operatives and smaller percentages of managers and sales workers.

While the total usable sample included 651 respondents, not all are included in all analyses. Some analyses, for example, had to be begun before the data from all five organizations became available. Some analyses exclude specified population categories, or exclude individuals lacking a full set of the required variables. These deviations from the full sample are specified in each chapter where relevant.

Table 1

Demographic and Occupational Distribution of Effectiveness in Work Roles Sample and Comparison National Survey Sample

Demographic or Occupational Characteristics	Percentage Distribution	
	Effectiveness in work roles sample (N=651)	Comparison national sample (N=2157)[a]
Sex		
Men	51.0%	62.1%
Women	49.0	37.9
Age		
Under 22	10.7	8.1
22-29	32.5	27.1
30-44	33.2	30.6
45-54	16.2	20.6
55 or older	7.4	13.6
Education		
Some grade school or less	2.9	4.7
Completed grade school	5.9	6.6
Some high school	18.2	14.2
Completed high school	41.6	38.4
Some college	19.9	20.9
College degree or more	11.5	15.2
Race[b]		
White	80.2	91.5
Black	19.8	8.5
Marital Status		
Never married	20.0	15.6
Married	68.5	74.7
Widowed, separated, divorced	11.5	9.7
Major occupation group[c]		
Professional and technical	14.3	15.6
Managers, officials, proprietors	9.8	15.4
Clerical	15.1	17.6
Sales	0.2	5.4
Craftsworkers, foremen	15.8	12.4
Operatives	29.0	18.5
Service workers, excluding private household	12.6	11.5
Laborers	3.1	3.5

---

[a] Source: Quinn, R., & Shepard, L. The 1972-73 Quality of Employment Survey. Ann Arbor: Institute for Social Research, 1974.

[b] Excludes minority races other than blacks.

[c] Based on 1960 Census codes excluding farmers and private household workers.

Measures

The Interview

The Survey Research Center and the respective employing organizations sent advance letters to all employees in the sampled departments. Since the hospital and one of the automotive supply organizations were unionized, union members were sent an additional letter from their union leaders. These letters outlined the general purposes of the study and requested cooperation. Appointments with the employees included in the sample were arranged by the professional interviewers who were in each case responsible for the interview. The interviews were conducted in most cases in the respondents' homes.

In order to test the clarity and appropriateness of the questions, workers from other organizations (not those studied) were interviewed prior to assembling the final version of the interview. Several pretest interviews were conducted by each of the professional interviewers who were to interview the respondents in the sample. The research staff met with the interviewers subsequent to these interviews, and decisions were made to change just a few questions in order to make them clearer or to omit them from the interview.

Formats Used in the Interview. The interview is reproduced in full in Volume II, Appendix A. The questions used five different formats.

1. Oral questions. Questions were asked orally by the interviewer who wrote down each person's verbatim response in the appropriate location (e.g., questions 3 and 4).

2. Show cards. These were cards with a set of printed response categories appropriate to several of the questions in the interview. The

interviewer asked the question, and the respondent was asked to choose among the alternative response categories (e.g., question 2).

3. Card sorts. These included two differently colored sets of four-inch by six-inch cards. Each card in the first set had one question printed on it, while the second smaller set had a response category on each card. The second set, the response categories, were laid out before the respondent who then sorted the question cards into piles next to the appropriate response categories (e.g., question 25).

4. Question booklet. This consisted of a few multiple-choice questions which the respondent read and then answered by marking the appropriate response categories. The question booklet is reproduced in Volume II, Appendix C.

5. Supervisor rating forms. Those respondents who stated that they supervised others as a part of their jobs were given one-page rating forms which they were requested to fill out for each of their subordinates. These subordinates (most of whom were also respondents in the study) were rated on eight dimensions: quality of work, quantity of work, creativity, lateness, dependability, liking for work, liking for responsibility, and getting along with co-workers. These forms, along with pre-stamped envelopes addressed to the Survey Research Center, were left with the supervisors for subsequent mailing. The supervisor rating form is reproduced in Volume II, Appendix E.

Content of the Interview. A considerable range of content areas were covered in the interviews.

1. The job. The respondents were asked three types of questions regarding their jobs:

a. Importance: Questions assessing the importance of various aspects of the job to the respondent constituted the first set of job-related questions.

b. Description: The second, and major, set of job-related questions was concerned with objective descriptions of working conditions. However, the degree of objectivity that was possible varied considerably; the questions ranged from such objective matters as annual income and fringe benefits to more judgmental matters such as the degree of creativity required by the job. The validity of the resulting Quality of Employment index is discussed in Chapter 3. Also included were questions regarding possible dangerous conditions at work, the number of work-related injuries and accidents sustained by the respondent, and the social conditions at work. In addition to these descriptive questions regarding the job, two further types of questions were asked. The first of these concerned the degree of financial equity that the respondent perceived himself or herself as having e.g., whether the respondent thought his or her pay to be fair compared to other people working in the same organization and also those working for a different employer. The second set concerned the degree to which the respondent perceived various rewards (and punishments) as being contingent on his or her performance, e.g., if he or she did a good job, would he or she get a promotion.

c. Attitudes: Among the job-related attitudes measured in the interview were job satisfaction, both overall and with respect to various specific aspects of the job, and motivation, including the amount of effort expended to perform well. Two other job-related areas that were tapped were depression and self-esteem felt in connection with the job.

2. Physical and mental health. In addition to the measures of job-related depression and self-esteem mentioned above, the interview included several other measures assessing the physical and mental health of the respondent (e.g., a list of psychosomatic symptoms).

3. Social and political involvement. A subset of the questions in the interview dealt with the kinds of social activity in which the respondents engaged, their leisure activities, the organizations to which they belonged, and their levels of political activity.

4. Personality and other characteristics. The interview contained selected questions (as opposed to complete scales) to be used in indices of certain personality characteristics of the respondent, e.g., rigidity, need for affiliation, need for approval, tolerance of ambiguity, and internal-external locus of control. Also included were questions concerning demographic characteristics of the respondent.

5. Interviewer's observations. At the conclusion of the interview the interviewer noted some physical characterisitcs of the respondent which would have been easily observed during the interview. These included the respondent's race, sex, weight, height, speech defects, and physical disfigurement.

The time referent for most of the questions in the interview was the present, i.e., respondents were asked to describe their jobs as they were at the time. For a small subset of the questions that dealt with rare events, e.g., garnishment of wages, a longer and retrospective time frame was used.

The average length of the interview was approximately 90 minutes.

Personnel Records

Data were obtained directly from the personnel records of the five organizations. The organizations varied in the kinds of information included in their records, as well as the completeness and orderliness of that information. The data that were obtained, therefore, varied somewhat from one organization to the next, though an attempt was made to match the information across organizations where possible. A list of the types of information gathered from each of the five organizations is shown in Table 2.

On-the-job Observations

The third set of data was obtained through on-the-job observations. This technique essentially consisted of having trained observers watch the respondent at work, and describe the respondent, as well as his or her coworkers, supervisor, and work environment.

Format of the Observation Booklet. As in the interview, the Observation Booklet (Appendix G) used various formats for recording the observations.

1. Physical descriptions. These were descriptive statements which each observer checked as being true or false of the work environment. Each indicated the presence or absence of something in that environment.

2. Job descriptions. The observer recorded how true each of a series of statements was with respect to the job and work environment. The difference between these and the previous set of statements was that more judgment was required of the observer in recording a scaled response to a statement

Table 2

Information Collected from Each of the Five Organizations' Personnel Records[a]

Information	Hospital	Printer	First Automotive Supplier	Second Automotive Supplier	R&D Firm
Pay	X	X	X	X	X
Employment date	X	X	X	X	X
Seniority-present job	X	X	X	0	0
Promotions received	X	X	X	0	0
Demotions received	X	X	X	0	0
Pay increases	X	X	X	0	X
Pay decreases	0	X	X	0	0
Supervisor's ratings	X	X	X	0	0
Commendation letters	X	X	X	0	0
Letters of reprimand	X	X	X	0	0
Education	X	X	X	X	X
Work-related injuries and illnesses	X	X	X	X	0
Union membership	X	0[b]	0[b]	X	0[b]
Number of jobs bid for	X	0	0	0	0
Age	X	X	X	X	X
Height	X	X	0	0	X
Number of hours worked per month	X	0	0	0	0
Shift	X	0	0	0	0
Vacation taken	X	0	X	X	X
Sick leave taken	X	X	X	X	X
Holidays taken	X	0	X	0	0
Absenteeism	X	X	X	X	X
Overtime worked	X	0	0	0	0
Lateral job transfers	0	X	X	0	0
Lateness	0	X	X	0	0
Termination	0	X	0	X	0
Garnishments	0	0	X	X	0
Hourly/Salaried	0	X	X	0	X

[a] X=this type of information was collected; 0=this type of information was not collected.

[b] There was no union.

such as "The work area is clean" (e.g., question A 2).

3. Anchored scales. These scales involved still more judgment on the part of the observer and took the form of seven-point scales referring to various aspects of the job. In order to reduce somewhat the difficulty of judgment, the scales included short descriptive statements at each of the two extremes and at the middle anchor (e.g., questions A3-A28).

4. Tallies. The observer also recorded the frequency of occurrence of various incidents, e.g., the frequency of interactions the subject had during the observation period, question D6.

Content of the Observations. The content of the observations was intended to parallel that of the interviews, although not inclusively. Those concepts that were tapped, however, were generally approached by the use of more than one observer rating format. Descriptions of the respondent's job, work group, supervisor, and working conditions constituted the main foci of the observations. These job and environmental descriptions varied considerably in the degree of "subjectivity" that was involved in observing and recording them.

Observation Procedures. Each observation period consisted of a one-hour observation of the respondent on the job by a trained observer. The beginning time of the observation was decided upon by the observer and the respondent, such that it was convenient for both and not too atypical of the respondent's regular work day.

Upon arriving at the work site, the observer introduced himself or herself to the respondent, assured the respondent that the records of the observed work behaviors would be confidential, and ascertained whether the ensuing work period would be a "typical" one. Since the respondent was necessarily aware of the observer anyway, it was decided that observer-respondent contact for these purposes would be beneficial. However, the observers were directed to keep these verbal contacts as brief and neutral as possible.

During the first 10 minutes of the observation period, the observer familiarized himself or herself with the respondent's job without recording anything. If it was not obvious what the respondent was doing, the observer was allowed to ask questions at the end of this period. During the next 30 minutes, the period of recorded observations, the observer counted frequencies of events, e.g., of interactions that the respondent had. The last 20 minutes of the hour were devoted to "general observations", during which the observer filled out the main body of the Observation Booklet. After the observation hour had been completed, the observer spent approximately 15 minutes editing -- checking to make sure that the booklet had been filled out correctly and completely, that all the information had been entered. The time schedule of the observation hour appears in Table 3.

Each respondent was observed for two hours, an hour each by two different observers at two different times. In addition, there were some "validation" observations. These involved two different observers watching a respondent at the same time and separately rating the respondent and his or her job. These "validation" observations were conducted for the purpose of assessing inter-rater agreement. In all, a total of 564 employees were observed.

A detailed description of the observation procedures, along with a report of reliability and validity of the resulting measures, appears in Chapters 4-6.

Table 3

Observation Schedule

Time Period	Activity	Part of Observation Booklet Used (see Appendix G)
10 minutes	Orientation to the job and the worker (observer observes job and makes no ratings)	--
30 minutes 30 minutes	Specific observations; observer counts and categorizes work chunks and interactions[a]	Work sheet for counting data for section D[a]
20 minutes	General observations and ratings	Section A
Actual observation ends; observer moves away from work location		
15 minutes	Editing:	
	- Transferring data from work sheet[a]	Section D
	- Responding to observer questions	Sections B, C
	- Filling in administrative information	Section E

---

[a] The work sheet used for this recording is shown in Volume II, Appendix G.

Chapter 2

DATA ACQUISITION AND FIELD METHODS - PHASE II

Chapter 2

DATA ACQUISITION AND FIELD METHODS - PHASE II

Since much of the methodology for Phase II duplicates procedures used for Phase I, this description of data acquisition and field methods for Phase II emphasizes the departures from the methods used in the earlier phase. At the outset, it should be noted that the data collection of Phase II took place in Fall of 1974, or 18-20 months after Phase I. Although there were no large and consistent differences between the two phases in the economic fortunes of the three organizations studied in both phases, idiosyncratic economic events such as a strike at one of the automotive firms during Phase II may have contributed to between-phase differences.

Sample

The Phase II design required a resurvey of employees from all the five organizations that had been included in Phase I of the study. One of the five organizations (the research and development firm) had, however, too few employees remaining from Phase I and was dropped as a result. The remaining four organizations were contacted to explore their willingness to participate in Phase II. One, the printer, refused whereas the other three agreed to participate in Phase II. In the hospital, however, only three of the original four departments were contacted--the fourth, the laboratory research section of internal medicine also had too few Phase I respondents to warrant inclusion in Phase II.

The final Phase II sample of organizations consisted, therefore, of the two automotive suppliers and three departments of the hospital. Extensive work was done at each of the three organizations to determine the sample of respondents. On the basis of information available from the records of the organizations, four types of respondents were identified. These were:

1. Phase I respondent was working in the same job with the same employer;
2. Phase I respondent was working in a different job with the same employer;
3. Phase I respondent had terminated work with the Phase I employer; and
4. New individual had replaced the terminated Phase I respondent.

Category 4 was particularly hard to define, since jobs and job titles were often changed when employees terminated, and frequently it was difficult to determine whether the person who was hired was actually filling the job that had been vacated. Occasionally, a decision on the issue of replacement was made on the basis of conversations with the supervisor of the terminated Phase I respondent. The final decision was made on a rigid definition of "same job"--that is, the replacement had to hold the same job title in the same department as the terminated respondent.

Response Rates

The final sample consisted of 182 potential respondents from the hospital, 109 from the first automotive supplier, and 121 from the second automotive supplier. The interview response rates were 67.2% at the hospital, 92.7% at the first automotive supplier, and 39.7% at the second automotive supplier.[1] The rate for the three establishments combined was 65.9%. The final analysis sample consisted of 272 employees, 123 from the hospital, 101 from the first automotive supplier, and 48 from the second automotive supplier. Of these, 38 (30.9%) were supervisors at the hospital in Phase I, 21 (21.1%) were supervisors at the first automotive supplier, and 2 (4.2%) were supervisors at the second automotive supplier.

Demographic and Occupational Constitution of the Sample. The demographic and occupational consitution of the Phase II sample is reported in Table 1. The table shows some differences in sample characteristics between the two phases. Specifically, the Phase II sample had higher proportions of workers who were male, old, less educated, black, formerly married, and in professional or managerial occupations.

Measures

The Interview

Measurement of Changes in Job Descriptions between Phase I and Phase II. Those respondents who had the same job with the same employer in Phase I and Phase II were asked a series of questions about whether there had been a major change for the better, or for the worse in their jobs, or if their jobs had stayed the same between Phases I and II. These questions were asked with respect to a series of job facets, e.g., hours, supervision, pay, etc. (Volume II, Appendix B, question 5). In addition, they

---

[1] The second automotive supplier had recently had a strike and this might account for the relatively low response rate.

Table 1

Demographic and Occupational Constitution of the Phase II Sample (N=272)

Sex	
Men	55.9%
Women	44.1
Age	
21 and under	3.3
22-29 years old	29.3
30-44 " "	36.7
45-54 " "	19.3
55 and older	11.5
Education	
Some grade school or less	5.2
Completed high school	8.5
Some high school	17.8
Completed high school	34.4
Some college	24.8
College degree or more	9.3
Race	
White	75.6
Black	24.4
Marital Status	
Never married	17.4
Married	67.4
Widowed, separated, divorced	15.2
Major Occupational Group	
Professional, technical & kindred	19.6
Managers, officials, etc. except farm	14.8
Clerical & kindred	10.0
Sales workers	--
Craftsmen, foremen, etc.	6.6
Operatives & kindred	32.5
Service workers, except private households	13.7
Laborers except farm & mine	3.0

were asked about their overall reaction to the job, and whether changes could be attributed to work versus other causes (Volume II, Appendix B, question 6).

Respondents who did not have the same job in Phase I but were nevertheless working at that time were also asked a series of questions about the differences between their jobs at Phase I and Phase II (Volume II, Appendix B, questions 13 and 14).

Job Descriptions. The roster of items that tapped descriptions of the respondents' jobs was expanded to include facets of the job that had not been measured in Phase I (Volume II, Appendix B, question 33).

Criterion Measures. A new list of variables measuring the employees' responses to their jobs was included. This list included such variables as self-reports of performance, job involvement, and reactions to co-workers and direct supervisors (Volume II, Appendix B, question 12).

Personality Measures. The roster of personality measures tapped was reduced so that only three concepts were assessed: need for affiliation, rigidity and tolerance of ambiguity (Volume II, Appendix B, question 121).

Supervisor rating form. While the content of the supervisor rating form stayed the same, the procedure for its administration was changed owing to problems encountered in Phase I. In Phase II, respondents were asked the name of their supervisor. Each supervisor so mentioned was mailed a rating form for each subordinate mentioning him or her as the supervisor. Supervisor ratings were requested and obtained only for these respondents, and supervisors were asked to mail back the completed rating forms to the Survey Research Center.

Deletions. Certain concepts tapped in Phase I were not included in Phase II because analysis of the Phase I data revealed problems with the scales or redundancy among concepts. These deletions included such variables as pace control, equity, parents' education, and internal versus external locus of control. In addition, some items within indices were deleted.

Response Alternatives. The number of response alternatives specified for particular questions was occasionally changed. For example, personality variables using a true/false dichotomy in Phase I were changed to four-point scales from "very true" to "not at all true".

Personnel Records

The information collected from the personnel records of the three organizations differed somewhat from the comparable information gathered in Phase I. Some information (e.g., absences, termination) was collected for each of the months between the two interviews; other information (e.g., overtime, shift) was collected for only the two months preceding the Phase II interview. A list of the types of information gathered from each of the three organizations appears in Table 2. It should be pointed out that an attempt was made to collect information on these variables for all employees interviewed in Phase I rather than only those interviewed in Phase II.

On-the-job Observations

As reported in detail in Chapter 7, a number of changes were made in the methodology for obtaining on-the-job observations. More time was devoted during training to explaining to observers the concepts being measured by their ratings; agreement with expert ratings replaced agreement among observers as the major criterion of observer selection;

Table 2

Information Collected from Each of the Three Organizations' Personnel Records (Phase II)[a]

Information	Hospital	First Automotive Supplier	Second Automotive Supplier
Seniority in present job	O	X	X
Work-related illnesses or injuries	O	X	X
Union membership	O	O[b]	X
Number of hours worked per month	X	X	O
Overtime worked	X	O	O
Shift	O	X	X
Vacation taken	X	X	X
Holidays taken	X	X	X
Sickleave taken	X	X	O
Jury duty	O	X	O
Enforced layoff due to strike	O	O	X
Discipline	O	X	O
Other excused and unexcused absences	X	X	X
Termination	X	X	X

---

[a] X=this type of information was collected; O=this type of information was not collected.

[b] There was no union.

new videotapes encompassing a wider range of job types were developed to facilitate the training of observers; unnecessary or problematic items were eliminated from the observation booklet; questions were added to help the observer focus attention at the outset on critical elements of the job; the series of general observations was scheduled before rather than after the sequence of specific observations (Table 3); and questions regarding interpersonal activities were also included.

As regards the sample of respondents, the observations were conducted in only two of the three organizations studied in Phase II. The third organization, the first automotive supplier, was not observed for three reasons: tension in the site owing to layoffs of personnel; extensive changes in employees' jobs following Phase II interviews owing to layoffs and recalls; and dwindling numbers of available and trained observers. In all, 147 respondents were observed in the two organizations. Of these, 100 were observed for a single one-hour observation period. The remaining 47 were observed by three observers, once by one observer alone, and once by two different observers watching simultaneously. This design was implemented for validation purposes.

Telephone Interview

Respondents from Phase I who had left their original jobs and employers by Phase II were briefly interviewed by telephone. They were asked why they had left their Phase I jobs and, in the case of those who had quit voluntarily, whether their new jobs were better or worse than their old ones on a list of job facets.

Table 3

Suggested Time Schedule for Observations

Part of observation booklet	Activity	Time allotted for each part	
		Minimum	Maximum
I	Introduction to employee and orientation to job (no ratings)	5 min	15 min
II	General observation (no rating)	15 min	45 min
III	Structured Observation of the job	--	15 min
IV	Rating the job	--	15 min
V	Administrative Information	--	--
VI	Editing (done away from the job)	--	--

Chapter 3

MEASURING EFFECTIVENESS IN WORK ROLES FROM THREE PERSPECTIVES

by

Cortlandt Cammann
Robert P. Quinn
Terry A. Beehr
Nina Gupta

ABSTRACT

This paper introduces three perspectives on the effectiveness of work roles: specifically, those of the employer, the worker, and society at large. It outlines the quality-effectiveness strategy in which working conditions that bear a relationship to criteria of effectiveness from any of the three perspectives are said to indicate high quality of employment. Extensive listings of working conditions and criteria of effectiveness were reduced to 31 and 11 items, respectively, with the 31 items on working conditions subsequently combined into an overall index of Quality of Employment. The pattern of relationships among the criteria of effectiveness was determined and the degree of variation of this pattern across different subgroups was ascertained. Efforts to assess the convergent validity of various measures of working conditions drew upon three modes of data collection: personal interviews, standardized observations, and company records. Analysis revealed a substantial degree of convergent validity across both the different modes of measurement and different workers in roughly the same jobs (consensual validation). An examination of construct validity established, as expected, that the Quality of Employment Indicator related more closely to job satisfaction than to the other criteria of effectiveness. The relationships between the Quality of Employment Indicator and the criteria of effectiveness were successfully duplicated in most subsamples of workers. Finally, both methodological and conceptual strategies were noted for improving the predictability of effectiveness from the Quality of Employment Indicator. Several attempts at refinement via methodological strategies yielded only marginal gains.

Chapter 3

MEASURING EFFECTIVENESS IN WORK ROLES FROM THREE PERSPECTIVES

PRIORITIES AND PERSPECTIVES

If priorities are to be assigned among various policies and programs intended to improve working conditions, the development and execution of these policies and programs will obviously be very costly. To justify the cost of any such policy or program it must be anticipated that there is some resulting benefit. But who should be the beneficiary? Priorities should clearly be assigned to improving quality of working life with respect to those job facets that heighten the effectiveness of each worker. But from whose perspective should such effectiveness be judged? There are at least three distinct perspectives for evaluating effectiveness.

The first perspective, that of employers, assigns priorities in terms more relevant to the needs of those doing the employing rather than to the needs of those who are employed. As a result many previous investigations of the determinants of employee effectiveness, not to mention employers' normal business and accounting records, emphasize productivity, quality of output, absenteeism, turnover, and similar indicators of effectiveness.

A second perspective for assessing effectiveness is that of the worker. Workers form conclusions or expectations about their effectiveness in their work roles not only in terms of actual work performance, but also in more general terms of the costs and benefits associated with their work roles. Thus, a worker normally assesses his or her work role effectiveness in such familiar terms as earnings, access to promotion, avoidance of accident or discomfort, security, intrinsic satisfactions from the work done, and the like. The worker, too, engages in something like a cost-benefit calculation, with consequences that include such behavior changes as quitting, working harder, seeking improvement in rewards, seeking qualifications for a "better" job, and the like.

A third perspective can also be invoked: that of the community or the society. Some of the costs and benefits associated with work role effectiveness do not enter into the formal or informal accounting of either the employer or the worker. For example: underutilized skills are lost to the economy; the laid off worker drains the public budget; the economic impact of a work stoppage falls in the end upon the public; the income deficient worker burdens the society with a family prone to illness, future welfare costs, and substandard economic contribution.

Which of these three perspectives is the most appropriate one for resolving issues of costs and benefits in the allocation of priorities to policies and programs is not a question that can be resolved by research. It is fundamentally one of values. No amount of research can determine whether it is "better," for example, to have a large, docilely contented workforce that is under-producing to the degree that both workers and others in society resultingly suffer rather than to have a disgruntled, restless, unhealthy and even angry workforce whose behaviors nevertheless continue to raise the GNP, corporate profits, and the quality of their own lives as consumers. Decisions as to which of these two situations, or, for that matter, any situation that represents an imbalance of perspectives, is the "better" or "best" one are placed most appropriately in the laps of those individuals, groups, and organizations whose positions confer on them the legitimacy to make such decisions. Research can at best help such social planners to recognize which perspectives may at present be incompatible.

For purposes of priority setting, the importance of any aspect of a work role, therefore, ultimately depends upon the magnitude of its impact upon a variety of criteria of effectiveness, the differential significance of which is largely a matter of a somewhat arbitrary, but hopefully informed and humane, selection from among a possibly incompatible set of values. If future research is to establish causal relationships between working conditions and their effectiveness outcomes, upon which outcomes should this research focus? Employers would probably suggest that attention be paid first to those working conditions that directly affect the productivity of their workers. This view is, however, a very narrow one unless it can be demonstrated that once productivity is raised or maintained all other desired outcomes will follow. More defensible is a broader-based investigative attack upon determining the consequences of working conditions that are valued according to not one but all three of the perspectives described above.

THE QUALITY-EFFECTIVENESS STRATEGY

Early in 1969 the U.S. Department of Labor, in collaboration with The University of Michigan's Survey Research Center, instituted a survey research program designed to assess some of the conditions under which Americans work. It was hoped that the data, based on personal interviews with workers, would aid policymakers in evaluating the needs and problems of workers. The investigators defined "working conditions" in terms sufficiently broad to encompass not only existing labor standards areas (for example, wages, hours, health and safety, discrimination), but also such "new" areas as the content of workers' jobs and their supervision. The first report based on this study appeared in the Monthly Labor Review in 1971 (Herrick & Quinn, 1971). A source book of univariate and bivariate statistics was published the same year (Quinn, Seashore, Kahn, Mangione, Campbell, Staines, & McCullough, 1971), and many other papers and reports have appeared since.[1]

A critical decision had to be made two years later when the Department of Labor expressed an interest in repeating the national survey. While repetition was certainly useful for the identification of trends, mere repetition was not of much interest to any of the parties concerned. The problem boiled down to identifying those materials that would be repeated, those that would be added, and those that would be scrapped.

The "quality-effectiveness strategy" was evolved in order to resolve this problem[2] and to capitalize on the national survey's analytic (rather than descriptive) potentials. This strategy defines three general concepts:

1. Working conditions. The term "working conditions" refers to descriptions of characteristics of a worker's job obtained from any

---

[1] For a full list, see Appendix B to The 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974).

[2] Cross-cutting the application of this strategy was a continuing interest in 19 types of working conditions subsumed under the general heading of "labor standards problems" (Herrick & Quinn, 1971).

informed source. These descriptions may focus on any characteristic of the job from the cleanliness of the physical work environment to the degree of time pressure for performance, or from the degree of challenge the job provides to the type of supervision the worker receives. Since working conditions characterize the job, they are, therefore, independent of the individual who does the job. This means that they do not include the worker's evaluation of the conditions measured and that different people doing the same job should describe it similarly.

2. Effectiveness. The term "effectiveness" refers to states or events that have a positive or negative value to some person or set of people. Three such perspectives were described above: those of employees, their employers, and society as a whole.

3. Quality of employment. The term "quality of employment" refers to a judgment about working conditions based on the impact that the working conditions have on effectiveness judged from some perspective. Thus, good quality of employment from the perspective of a worker would be some combination of working conditions that produce health (one criterion of effectiveness as judged by the worker). Good quality of employment from the perspective of an employer would be working conditions that lead to a productive, profitable organization.

These definitions provided a standard for selecting measures of working conditions to be carried forward into the second national survey: select only those that had a demonstrable association with some criterion of effectiveness, that is, select indicators of quality of employment. Over 30 indicators were thus selected empirically and constituted the Quality of Employment index (Barnowe, Mangione & Quinn, 1972). These indicators are listed in Table 1. The criterion used in that particular selection was, however, very limited--job satisfaction. It not only ignores the perspectives of employers and society at large but does not even begin to cover the outcomes of concern to employees. This limitation was overcome somewhat in the second national survey (Quinn & Shepard, 1974). Job satisfaction, while measured once again, plays a small part in the completed and projected analyses of that survey's data.

Far greater attention is being paid to two other criteria of effectiveness as judged from a worker's perspective--physical health and alcohol consumption.

Table 1

Components of Quality of Employment Indicator

Challenge

Worker's supervisor encouraged new ways of working

Worker's job required high level of skill

Worker's job allowed freedom as to how to do his or her work

Worker's job did not prevent him or her from using skills he or she would like to be using

Worker's supervisor let his or her subordinates alone unless they asked for help

Worker's job required learning new things

Worker's job required that he or she be creative

Worker's job involved doing a variety of things

Worker had exactly the education his or her job required

Worker's job allowed him or her to make a lot of decisions on his or her own

Worker had enough authority to tell others what to do

Worker's employer made available to him or her a training program for improving his or her skills

Resource Adequacy

Worker's supervisor maintained high standards in his or her work

Worker's supervisor knew his or her own job well

Worker had enough help from others with whom he or she worked

Worker had enough machinery and equipment to do his or her job well

Worker had enough facts and information to do his or her job well

Comfort

Worker had no problems with hours, work schedule, or with working overtime

Worker did not experience dangerous or unhealthy conditions on his or her job

Worker had enough time to do what others expected of him or her

The physical conditions of worker's job were pleasant and comfortable

(continued)

Table 1 (continued)

Comfort (continued)

- Worker had no problems with transportation to and from work
- Worker mostly determined whether he or she would work overtime on his or her job
- Worker did not work excessive hours
- Worker's supervisor did not insist that those under him or her work hard
- Worker did not have to take much time to get to work

Financial Rewards

- Worker's employer made many fringe benefits available to him or her
- Worker desired no additional fringe benefits
- Worker was a full time worker who received a high income from his or her job
- It was unlikely that worker's job would be automated
- It would be easy for worker to find a new job as good as his or her present one

EFFECTIVENESS IN WORK ROLES

Even an increased emphasis on the latter two criteria left a lot of ground uncovered. The Effectiveness in Work Roles study was conceived as one means of covering this ground. While its goals were principally those of validation, the latter term includes a lot of territory when it encompasses the notion of construct validation. Moreover, the quality-effectiveness strategy makes the validation of quality of employment indicators largely a matter of testing propositions that relate working conditions to criteria of effectiveness. The quality-effectiveness strategy therefore transforms many issues of "validation" into "substantive" issues and vice-versa.

Four of the major goals of the Effectiveness in Work Roles study were:

1. to develop improved conceptual models for assessing the nature of working conditions that are associated with effectiveness in the performance of work roles as viewed from the value perspectives of the employee, the employer, and society;
2. to examine the structure of effectiveness measures;
3. to assess the correspondence among measures of working conditions obtained through different techniques; and
4. to determine the relationships between measures of quality of employment and measures of effectiveness in work roles.

Initially many criteria of effectiveness were obtained. These were ultimately reduced to a set of eleven, embodying all three perspectives. Moreover, the data contained measures, often from multiple sources, of approximately 50 distinct types of working conditions. Lest the analysis become a "fishing expedition" among over 500 associations between working conditions and criteria of work-role effectiveness, the decision was made to limit drastically the number of working conditions attended to. For this reason, the entire analysis focused only upon the Quality of Employment Indicator, the components of which were listed in Table 1. These components were further grouped into four more general areas--Comfort, Challenge, Resource Adequacy, and Financial Rewards (see sub-headings in Table 1).

With these restrictions, several questions were asked:

1. What is the structure of criteria of effectiveness in work roles? This question is answered in the "Measuring Effectiveness from Three Perspectives" below

2. How much agreement is there among measures of working conditions obtained from different sources? More specifically, how much agreement is there among measures of quality of employment obtained through personal interviews, on-the-job observations, and employers' records? These questions are answered in the "Convergent Validity" section

3. How does quality of employment relate to various criteria of effectiveness? This is answered in the section dealing with "Construct Validity"

4. How can the Quality of Employment Indicator be improved? The answer to this will be found in the final section.

METHOD

Sample

The three establishments included in the sample were a hospital, a printing company, and a manufacturer of automobile accessories--all in the midwest. People working less than 20 hours per week were excluded. Certain departments in each establishment, rather than the entire establishment, were used as the units from which the sample was drawn. Within each of the departments selected, all supervisory personnel were included in the sample; non-supervisory employees were systematically sampled at lower rates.[1] Subsequently, those employees who felt that they had not been on their present jobs long enough to answer the job-related questions adequately were screened out by the interviewers.

The departments selected in the hospital were those in charge of patients' nutrition, general services (custodians, elevator operators, ambulance drivers, etc.), respiratory therapy, and the laboratory research section of the internal medicine department. Because the first two departments were quite large, 50 percent of non-supervisors and 100 percent of the supervisors were included.[2] This resulted in a total of 54 supervisors and 80 non-supervisors in the patients nutrition department and 17 supervisors and 61 non-supervisors in the general services department.

---

[1] All supervisors were included in the sample because performance ratings of each respondent by his or her supervisor were needed. This was one way to insure that each respondent's supervisor was in the sample. Information regarding the supervisory status of the employee was provided by the organizations studied.

[2] The sample was drawn by choosing every second employee from a list furnished by the hospital. Whether the first or second individual on the list was the first respondent chosen was determined randomly.

All supervisors and non-supervisors from the respiratory therapy and the laboratory research section of the internal medicine department were included in the sample, primarily because these departments were relatively small in number--54 employees in the first and 13 in the second department. Doctors and nurses were excluded from the sample because of the problems anticipated in attempting on-the-job observations of their behaviors. Technical personnel, respiratory therapists, and laboratory technicians were all included in the sample.

At the printing company all departments were sampled with two exceptions: the top level managers (who also owned the firm); and the sales department, which was spread all over the country. Four small departments--maintenance, finance, personnel, and engineering--were sampled at a rate of 100 percent among both supervisors and non-supervisors. All the supervisors, and 66.7 percent of the non-supervisors were sampled[1] from the remaining departments, viz., sales services, warehouses, and eight departments dealing directly with various stages of production. Thus, the sample for the printing company included all 42 supervisors and 177 of 272 non-supervisors in the departments selected for study.

The sample at the automobile supply organization consisted mainly of the basic production departments. The major exclusions were the sales, research and development, and accounting units. Within the sampled units, 100 percent of the supervisors and 50 percent of the non-supervisors were included in the sample. The resulting sample included 30 supervisors and 138 non-supervisors.

Response rates. The interview response rates were 76.9 percent at the hospital, 79.0 percent at the printer and 71.4 percent at the automobile supplier. The rate for the three establishments combined was 76.1 percent. Seven and six-tenths percent of the sampled employees refused to be interviewed, and 16.3 percent were not interviewed for other reasons. The final analysis sample contained 506 employees.

Demographic and occupational constitution of the sample. The demographic and occupational constitution of the sample is reported in Table 2, together with comparative statistics from a 1973 national sample

---

[1] The sample was systematic, e.g., 110110110, etc.

Table 2

Demographic and Occupational Distribution of Effectiveness in Work Roles Sample and Comparison National Survey Sample

Demographic or Occupational Characteristic	Percentage Distribution	
	Effectiveness in Work Roles sample (N = 506)	Comparison national sample (N = 2157)[a]
Sex		
Men	46.4%	62.1%
Women	53.6	37.9
Age		
21 or under	8.7	8.1
22-29	31.2	27.1
30-44	35.8	30.6
45-54	15.6	20.6
55 or older	7.9	13.6
Education		
Some grade school or less	2.2	4.7
Completed grade school	4.9	6.6
Some high school	15.0	14.2
Completed high school	42.5	38.4
Some college	22.7	20.9
College degree or more	12.6	15.2
Race[b]		
White	77.1	91.5
Black	22.9	8.5
Marital Status		
Never married	21.9	15.6
Married	67.2	74.7
Widowed, separated, or divorced	10.7	9.7
Major Occupation Group[c]		
Professional and technical	17.0	15.6
Managers, officials, and proprietors	10.7	15.4
Clerical	18.8	17.6
Sales	0.2	5.4
Craftsworkers and foremen	15.6	12.4
Operatives	19.4	18.5
Service workers, excluding private household	16.2	11.5
Laborers	2.0	3.5

[a] 1972-73 Quality of Employment Survey.

[b] Excludes minority races other than blacks.

[c] Based on 1960 Census codes. The Effectiveness in Work Roles sample contained no farmers or private household workers, and workers in these occupations have been excluded from the bases of the comparison national statistics.

of workers (Quinn & Shepard, 1974). Since the sample was not intended to be representative of the national population, the differences between it and the national sample are not surprising, but it may be important to keep these differences in mind when the results of the study are interpreted. Demographically, the sample had a higher proportion of blacks, women, single people, and young people than the comparison national sample. Since the sample was drawn from three organizations performing rather specific operations, the occupational subgroups represented in it could be expected to be quite different from those in the national sample. Given this expectation, the percentages in each category for the two samples were remarkably similar. The major differences were that the Effectiveness in Work Roles' sample included a larger percentage of service workers and smaller percentages of managers and sales workers.

Measures: The interview

The Survey Research Center and the employing organizations sent letters to all employees in the three organizations.[1] Union members received an additional letter from their union leaders. These letters outlined the general purposes of the study and requested workers' cooperation. Appointments for interviews with the employees sampled were arranged by the professional interviewers who were responsible for the interviews. The interviews were usually conducted in the respondents' homes.

In order to test the clarity and appropriateness of the questions, workers not in the organizations studied were interviewed prior to assembling the final version of the interview. A pretest was conducted by the same professional interviewers who were to interview the respondents in the sample. The research staff met with the interviewers subsequent to these interviews, and decisions were made either to change questions in order to make them clearer or to omit them from the interview.

---

[1] In the case of the hospital, letters were sent only to the employees of the four departments sampled.

Table 2

Demographic and Occupational Distribution of Effectiveness in Work Roles Sample and Comparison National Survey Sample

Demographic or Occupational Characteristic	Percentage Distribution: Effectiveness in Work Roles sample (N = 506)	Percentage Distribution: Comparison national sample (N = 2157)[a]
Sex		
Men	46.4%	62.1%
Women	53.6	37.9
Age		
21 or under	8.7	8.1
22-29	31.2	27.1
30-44	35.8	30.6
45-54	15.6	20.6
55 or older	7.9	13.6
Education		
Some grade school or less	2.2	4.7
Completed grade school	4.9	6.6
Some high school	15.0	14.2
Completed high school	42.5	38.4
Some college	22.7	20.9
College degree or more	12.6	15.2
Race[b]		
White	77.1	91.5
Black	22.9	8.5
Marital Status		
Never married	21.9	15.6
Married	67.2	74.7
Widowed, separated, or divorced	10.7	9.7
Major Occupation Group[c]		
Professional and technical	17.0	15.6
Managers, officials, and proprietors	10.7	15.4
Clerical	18.8	17.6
Sales	0.2	5.4
Craftsworkers and foremen	15.6	12.4
Operatives	19.4	18.5
Service workers, excluding private household	16.2	11.5
Laborers	2.0	3.5

[a] 1972-73 Quality of Employment Survey.

[b] Excludes minority races other than blacks.

[c] Based on 1960 Census codes. The Effectiveness in Work Roles sample contained no farmers or private household workers, and workers in these occupations have been excluded from the bases of the comparison national statistics.

of workers (Quinn & Shepard, 1974). Since the sample was not intended to be representative of the national population, the differences between it and the national sample are not surprising, but it may be important to keep these differences in mind when the results of the study are interpreted. Demographically, the sample had a higher proportion of blacks, women, single people, and young people than the comparison national sample. Since the sample was drawn from three organizations performing rather specific operations, the occupational subgroups represented in it could be expected to be quite different from those in the national sample. Given this expectation, the percentages in each category for the two samples were remarkably similar. The major differences were that the Effectiveness in Work Roles' sample included a larger percentage of service workers and smaller percentages of managers and sales workers.

Measures: The interview

The Survey Research Center and the employing organizations sent letters to all employees in the three organizations.[1] Union members received an additional letter from their union leaders. These letters outlined the general purposes of the study and requested workers' cooperation. Appointments for interviews with the employees sampled were arranged by the professional interviewers who were responsible for the interviews. The interviews were usually conducted in the respondents' homes.

In order to test the clarity and appropriateness of the questions, workers not in the organizations studied were interviewed prior to assembling the final version of the interview. A pretest was conducted by the same professional interviewers who were to interview the respondents in the sample. The research staff met with the interviewers subsequent to these interviews, and decisions were made either to change questions in order to make them clearer or to omit them from the interview.

---

[1] In the case of the hospital, letters were sent only to the employees of the four departments sampled.

The formats used in the interview. The interview is reproduced in full in Appendix A.[1] The questions used five different formats:

1. Questions asked orally by the interviewer who recorded each person's verbatim response in the appropriate location (e.g., questions 3 and 4 in Appendix A).

2. Show cards: These were cards with a set of printed response categories which referred to several of the questions in the interview. The interviewer asked the question, and the respondent read aloud the category he or she chose. These show cards were used to help the respondents to remember the alternative response categories (e.g., question 2 in Appendix A).

3. Card sorts: These included two differently colored sets of four-inch by six-inch cards. Each card in the first set had one question printed on it, while the second set had a response category on each card. The second set, the response categories, were laid out before the respondent who then sorted each of the questions from the first set of cards into piles next to the appropriate response categories (e.g., question 25 in Appendix A).

4. Question booklet: This consisted of a few multiple-choice questions which the respondent read and then answered by marking the appropriate response categories. The question booklet is reproduced in Appendix C.

5. Supervisor rating forms: Those respondents who stated that they supervised others as part of their jobs were given one-page rating forms which they were requested to fill out for each of their subordinates. These subordinates (most of whom were also respondents in the study) were rated on eight dimensions: quality of work, quantity of work, creativity, lateness, dependability, liking for work, liking for responsibility, and getting along with co-workers. These forms, along with prestamped envelopes addressed to the Survey Research Center, were left with the supervisors for subsequent mailing. The supervisor rating form is reproduced in Appendix E.

---

[1] With one exception, all references in the text of this chapter to appendices refer to methodological appendices in Volume II. The exception is an appendix specific to this chapter which appears at the end of this chapter and which concerns the measurement of dangerous and unhealthy conditions.

The content of the interview. A considerable variety of content areas was covered in the interviews.

1. The job: The respondents were asked three types of questions regarding their jobs:

   a. Importance: Questions assessing the importance of various aspects of the job to the respondent constituted the first set of job-related questions.

   b. Description: The second, and major, set of job-related questions was concerned with objective descriptions of working conditions. However, the degree of objectivity that was possible varied considerably; the questions ranged from such objective matters as annual income and fringe benefits to more subjective matters such as the degree of creativity required by the job. Also included were questions regarding possible dangerous conditions at work, the number of work-related injuries and accidents sustained by the respondent, and the social conditions at work. In addition to these descriptive questions regarding the job, two further types of questions were asked. The first of these concerned the degree of financial equity that the respondent perceived himself or herself as having, e.g., whether the respondent thought his or her pay to be fair compared to that of his or her colleagues. The second set concerned the degree to which the respondent perceived various rewards (and punishments) as being contingent on his or her performance, e.g., if he or she did a good job, would he or she get a promotion?

   c. Attitudes: Among the job-related attitudes tapped in the interview were job satisfaction, both overall and with respect to various specific aspects of the job, and motivation, including the amount of effort expended, to perform well. Two other job-related variables that were tapped were depression and self-esteem that the respondent felt in connection with the job.

2. Physical and mental health: In addition to the job-related depression and self-esteem mentioned above, the interview included several other measures assessing the physical and mental health of the respondent.

3. Social and political involvement: A subset of the questions in the interview dealt with the kinds of social activity in which the respondents engaged, the organizations to which they belonged, and their levels of political involvement.

4. Personality and other characteristics: The interview contained questions to be used in indices of certain personality characteristics of the respondent, e.g., rigidity, need for affiliation, need for approval, tolerance of ambiguity, and locus of control. Also included were questions concerning demographic characteristics of the respondent.

5. Interviewer's observations: At the conclusion of the interview, the interviewer noted some physical characteristics of the respondent which would have been easily observed during the interview. These included the respondent's race, sex, weight, height, speech defects, and physical disfigurement.

The time referent for most of the questions in the interview was the present, i.e., respondents were asked to describe their jobs as they currently were. For a small subset of the questions that dealt with rare events, e.g., garnishment of wages, a longer time frame was used.

Measures: Personnel records

A second source of data was personnel records from each of the three organizations. The organizations varied in the kinds of information that were included in their records, as well as the completeness and orderliness of that information. The data that were obtained, therefore, varied somewhat from one organization to the next, though an attempt was made to match the information across organizations where possible. A list of the types of information gathered from each of the three organizations is shown in Table 3.

Measures: On-the-job observations

The third set of data was obtained using a relatively new technique, on-the-job observations. This technique essentially consisted of having trained observers watch the respondent at work, and describe the respondent, as well as his or her co-workers, supervisor, and work environment.

Table 3

Information Collected from Each of the Three Organizations' Personnel Records[a]

Information	Collected from Hospital Records	Collected from Printer's Records	Collected from Automotive Supplier's Records
Pay	X	X	X
Employment date	X	X	X
Seniority in present job	X	X	X
Promotions received	X	X	X
Demotions received	X	X	X
Pay increases	X	X	X
Pay decreases	0	X	X
Supervisor's ratings	X	X	X
Letters of commendation	X	X	X
Letters of reprimand	X	X	X
Education	X	X	X
Work-related injuries and illnesses	X	X	X
Union membership	X	0[b]	0[b]
Number of jobs bid for	X	0	0
Age	X	X	X
Height	X	X	0
Number of hours worked per month	X	0	0
Shift	X	0	0
Vacation taken	X	0	X
Sick leave taken	X	X	X
Holidays taken	X	0	X
Absenteeism	X	X	X
Overtime worked	X	0	0
Lateral job transfers	0	X	X
Lateness	0	X	X
Termination	0	X	0
Garnishments	0	0	X

[a] X = this type of information was collected; 0 = this type of information was not collected.

[b] There was no union.

The format of the Observation Booklet. As in the interview, the Observation Booklet (Appendix G) used different formats for recording the observations.

1. Physical descriptions: These were purely descriptive statements which each observer checked as being true or false of the work environment. Each indicated the presence or absence of something in that environment.

2. Job descriptions: The observer recorded how true a series of statements were with respect to the job and work environment. The difference between these and the previous set of statements was that more judgment was required of the observer in recording a response to these, e.g., "the job is meaningful." See, for example, question 2 in Appendix G.

3. Anchored scales: These scales involved still more judgment on the part of the observer and took the form of seven-point scales referring to various aspects of the job. In order to reduce somewhat the degree of judgment that the observer would have to make, the scales included short descriptive statements at each of the two extremes and at the middle anchor. See, for example, questions A3-A28 in Appendix G.

4. Tallies: The observer also recorded the frequency of occurrence of various incidents, e.g., the frequency of interactions the subject had during the observation period. See, for example, question D6 in Appendix G.

The content of the observations. The content of the observations was similar to that of the interviews, although it was not as inclusive. Those concepts that were tapped, however, were generally approached by the use of more than one rating format. Descriptions of the respondent's job, work group, supervisor, and working conditions constituted the main foci of the observations. These job and environmental descriptions varied considerably in the degree of "subjectivity" that was involved in observing and recording them.

The observation procedures. Each observation period consisted of a one-hour observation of the respondent on the job by a trained observer. The beginning time of the observation was decided upon by the observer and the respondent, such that it was convenient for both and not too atypical of the respondent's regular work day.

Upon arriving at the work site, the observer introduced himself or herself to the respondent in order (1) to assure the respondent that the records of the observed work behaviors would be confidential and (2) to ascertain whether the ensuing work period would be a "typical" one. Since the respondent was necessarily aware of the observer anyway, it was decided that observer-respondent contact for these purposes would be beneficial. However, the observers were directed to keep these contacts short.

During the first 10 minutes of the observation period, the observer familiarized himself or herself with the respondent's job without recording anything. If it was not obvious what the respondent was doing, the observer was allowed to ask a question at the end of this period. During the next 30 minutes, the period of specific observations, the observer counted frequencies of events, e.g., of interactions that the respondent had. The last 20 minutes of the hour were devoted to "general observations," during which the observer filled out the main body of the Observation Booklet. After the observation hour had been completed, the observer spent approximately 15 minutes editing--checking,to make sure that the booklet had been filled out correctly and completely, that all the information had been transferred to it. The time schedule of the observation hour appears on the next page.

Each respondent was observed for two hours, an hour each by two different observers at two different times. In addition, there were some "validation" observations. These involved two different observers watching a respondent at the same time and separately rating the respondent and his or her job. These "validation" observations were conducted for the purpose of assessing inter-rater agreement. In all, there was a total of approximately 1,500 observation hours.

The Time Schedule of the Observation Hour

Time period	Activity	Part of Observation Booklet used (see Appendix D)
10 minutes	Orientation to the job and the worker (observer observes job and makes no ratings)	--
30 minutes	Specific observations; observer counts and categorizes work chunks and interactions[a]	Work sheet for counting data for section D[a]
20 minutes	General observations and ratings	Section A
Actual observation ends; observer moves away from work location		
15 minutes	Editing:	
	- Transferring data from work sheet[a]	Section D
	- Responding to observer questions	Sections B, C
	- Filling in administrative information	Section E

[a] The work sheet used for this recording is shown as the last page of Appendix G.

Training the observers

The training of the observers constituted a most crucial aspect of the use of the observation technique. Therefore, the study staff devoted considerable time and energy to ensuring that the observers were trained adequately, and that only "good" observers were selected.

Fifty-one "candidate observers" underwent the training. Most of them had had some college education, and at least a minimal background in psychology. They were recruited through notices posted at various places in The University of Michigan and an announcement in an undergraduate organizational psychology class.

One group of people observed at the hospital and the printing company, while a second group was employed at the automobile supplier, since the sites were situated at two geographically different locations. The group who observed at the printer and the hospital consisted of 36 candidates, from which 24 were finally selected. For the automotive supplier, eleven of 15 candidates were hired to make a total of 35 observers.

These groups were taught to rate jobs from video-tapes. In all, there were tapes of four jobs, although only three were used for the set of candidates at either location. Those three were chosen so as best to represent the types of jobs that would be observed at the respective locations.

The candidates observed and rated a video-taped job four times, with the first and third being the same tape. They used the same rating booklets that were to be used in the actual data collection. Between tapes they discussed their ratings in an attempt to improve their mutual agreement. The trainees were informed that such agreement on the fourth tape was the criterion for hiring and that all of them who were able to meet it would be hired.

Two criteria for the hiring decision were based on the deviations of the trainees' ratings on each item from the mean score of all trainees on the same item. If the shape of the distribution of a trainee's deviation scores was normal, and if the range of the distribution was relatively narrow, the trainee met the criteria. A normal distribution was interpreted as meaning that the deviations were random rather than systematic errors, and a narrow range meant that the trainee was in fact agreeing with the group of trainees as a whole.

The third criterion was based on the correlations between the scores of each trainee across all items and the scores of all other trainees across all items. High correlations indicated close agreement with other raters and, with the above criteria of shape and dispersion, led to a favorable hiring decision.

MEASURING EFFECTIVENESS FROM THREE PERSPECTIVES

If the Quality of Employment indicator shows that an employment situation is "good," this means that the employment situation should produce outcomes that are valued. In order to validate measures of quality of employment, it is therefore necessary to develop criteria that can be used to judge the value of a specific set of working conditions. One of the purposes of the Effectiveness in Work Roles study was to develop such criteria. The following pages will discuss the development of these outcome criteria for evaluating employment situations and will examine the relationships that exist among these criteria in order to show their implications for developing quality of employment indicators.

The first step in developing a set of criteria for evaluating different work situations was to identify the perspectives from which the criteria are valued. As the opening pages of this report indicated, three perspectives seemed especially critical: the perspective of the employee working in the situation, the perspective of his or her employer, and the perspective of society.

There are obviously a very large number of outcomes that can result from a work situation and that can be valued from each of these perspectives--more outcomes than can be usefully measured or related to quality of employment. As a result, only a small set of outcomes was chosen for measurement. These outcomes were felt to be the most important outcomes measurable within the scope of the study.

Outcomes from the perspective of the employee

Several outcomes valued from the perspective of the employee were identified. All of these were measured exclusively by workers' self-reports, because these criteria reflected the employees' own feelings, reactions, and physical states. Without using extensive observation and expensive medical examinations, there was no way of collecting the necessary information without asking employees about such matters. Whenever possible, the measures used were scales that had been shown to be reliable and valid in other investigations. These measures were the following:

Job satisfaction was measured by using the 1973 version of the job satisfaction measures developed by the Survey Research Center for use in the 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974). This measure of job satisfaction has two equally-weighted components: measures of the employee's satisfaction with specific facets of his or her job, and a measure of general satisfaction with the job. The reliabilities of these two components were .92 and .72, respectively.[1]

Depressed mood was also based on a scale used in the 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974). Its ten questions were drawn from the 20-question measure developed by Zung (1965). Its reliability was .77, and its correlations with the original 20-question measure were .95 among men and .95 among women.

Physiological health was measured by an instrument adapted from that of Belloc, Breslow, & Hochstim (1971).

Work-related illness and injury was measured by questions concerning both the frequency and severity of the injuries. The first question was, "Within the last year have you had any illnesses or injuries you think were caused or made more severe by any job you had during this period at (name of employer)?" If the answer was positive, the respondent was asked to list the illnesses or injuries and was asked if they had kept the respondent out of work for more than two weeks. The Illness and Injury scale was scored 1 if no illnesses or injuries were reported, 2 if there was only one and it had not kept the respondent from work for more than two weeks, 3 if there were two and they had not kept the respondent from work for more than two weeks, 4 if there were three illnesses and injuries and they had not kept the respondent from work for more than two weeks, and 5 if there was at least one illness or injury that had kept the respondent from work for more than two weeks.

---

[1] All reported reliabilities are Spearman-Brown internal consistency reliability estimates.

Life satisfaction was measured using the Overall Life Satisfaction scale used in the 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974). Its reliability was .88.

Self-esteem was measured by a three-item subscale of the self-esteem measure used in the 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974). Its items were seven-point semantic differential items with the anchors "successful/not successful," "doing my best/not doing my best," and "important/not important." The items referred to how respondents felt about themselves at work. The scale's reliability was .71.

Outcomes from the perspective of the employer

Five outcomes of working conditions were identified which would be valued by an employer.

Lateness was measured through self-reports. Employees were asked about the number of days they had been late in the last two weeks and whether they were late more often than other people at their place of work, about as often, or less often. These two measures were combined on the grounds that each measured slightly different aspects of lateness.

Absenteeism was obtained from employers' records. The measure was the number of absence episodes that had occurred in the month prior to the interview. An absence episode was defined as one wherein a person was absent on one or more days in succession. This particular measure was used because it minimized the contribution of both uncontrollable absences (such as those due to severe sickness) and long-term absences, the latter being more predictable and hence less costly to the employer.

Turnover tendency was measured by a question in the interview which asked the employees how likely it was that they would look for a job in the next year. It would, of course, have been preferable to get actual turnover figures, but this was not possible within the time span of this study. However, this particular question had been used successfully by Mangione (1973) to predict turnover.

Dependability and "job attitude" was measured by supervisors' ratings of the employees on three seven-point semantic differential scales: "very dependable/very undependable," "likes working very much/dislikes working very much," and "enjoys having responsibility/avoids having responsibility." Its reliability was .81.

Performance was measured by supervisory ratings of the quantity and quality of the employee's performance. The two measures were so highly related that they were combined into a single performance scale. The items were seven-point semantic differential items with the anchors "does very high quality work/does very low quality work," and "does a large amount of work/does very little work." Its reliability was .81.

Effort was measured by on-the-job observers who watched the employees at work. The observer rated the employee on three items. The first two were seven-point scales anchored at the ends and in the center:

To what extent does the employee work hard on his/her job?

1. Very little; he/she takes it easy on the job; he/she takes frequent breaks and spends much time not working; he/she does not put a great deal of energy into the job.
2.
3.
4. Moderate; he/she works moderately hard on his/her job.
5.
6.
7. Very much; he/she works very hard on the job; he/she takes very few breaks and always seems to be working; he/she puts a great deal of energy into the job.

To what extent is the employee efficient in doing his/her job?

1. Very inefficient; he/she does his/her job using a great deal of excessive motion; he/she works slowly; he/she does not appear to be very efficient.
2.
3.
4. Moderately efficient in doing his/her job.
5.
6.
7. Very efficient; he/she does his/her job using a minimum of motion.

The third item asked, "(How true is it that) the individual working on this job expends a lot of effort trying to perform his/her job well?"

The reliability of this scale on the occasion of a single observation was .88. Since employees were usually observed twice, an employee's score on the effort scale was usually an average of the scores from two observations.

Outcomes from the perspective of society

Many outcomes of value to employees and employers can also be viewed as important from a societal perspective. A society obviously values having productive and healthy members. However, there is also one outcome that may be regarded as primarily (but not exclusively) valued by the community, or society as a whole; viz., societal involvement. Three measures of societal involvement were used.

The political participation indicator was a mean of three questions. The first asked if the respondent had voted in the last national election. The second asked how interested the respondent had been in political campaigns during the pre-election period, and the third asked how often he or she had voted in past presidential elections.

The second measure of social participation assessed the involvement of the employee in formal organizations outside of work. The employee indicated in the interview whether he or she belonged to each of the following organizations: sports club or team; social or card playing group; church or synagogue; church or synagogue-connected group or association; lodge, fraternity, sorority or veterans association; labor union; a cooperative; national, ethnic, or racial association; professional association; social welfare or charity group; parents-teachers association; youth groups, e.g., being a Girl Scout leader or a Little League coach; country club; community center; discussion group; neighborhood or community improvement group; political club or organization.

The third measure of societal involvement assessed the degree to which the employee engaged in social recreation off the job. The respondent was asked how long ago he or she had participated in each of the following activities: went to the movies; went to a sports event; played in some sport yourself; ate in a restaurant; went to a bar or nightclub; went shopping for something besides groceries; went to a play or concert; went to a fair, exhibit, or museum; went to a class, talk, or lecture; got in touch with relatives, not counting those who live with you; got together with any friends other than your relatives--like going out together or visiting each others' homes; chatted on the telephone with friends; made a repair or fixed something around your home or apartment; sewed, embroidered or did other types of needle work; finished reading a book; read a whole article in a magazine; played cards or some other indoor game; worked on some hobby of yours; went hunting or fishing; met and talked with any people--other than those you met at work--that you had never met before; went out with your family; went for a pleasure trip in your car or drove somewhere where you could enjoy yourself. Respondents were also asked how often the previous summer they had gone camping or hiking, gone swimming or boating, gardened, or worked around the yard.

Refining the criteria of effectiveness

From the 16 valued outcomes described above, a reduced set of measures was derived to serve as criteria for evaluating working conditions because some of the criteria were conceptually and empirically related. The following set of measures are the result of the combining process.

Job satisfaction. The conceptual differences between facet-free job satisfaction and facet-specific job satisfaction did not necessitate maintaining these measures separately. As a result, each measure was standardized, and the two were combined to create a single job satisfaction scale (Quinn & Shepard, 1974).

Withdrawal. Absenteeism and lateness represent employee behaviors that indicate withdrawal from the organization. Since both types of behavior are costly to the organization, and since they represent different aspects of the same thing, it was decided to combine these variables into a single measure of withdrawal behavior. In order to create the withdrawal scale, both lateness and absenteeism frequency were converted into a common scale, and the scales were then averaged.

Performance. Both the supervisors' ratings of an employee's performance and the observers' ratings of the employee's effort were measures of the employee's overall contributions to the productivity of the organization. These measures were averaged to develop an overall performance measure.

Societal involvement. The three measures of social recreation, political participation, and number of organizations joined, are all aspects of societal involvement. As a result, these three scales were combined into a single measure. Furthermore, it was felt that while all three scales were descriptive of societal involvement, they were not all equally descriptive. The extent of a person's recreational activities was felt to be a better indication of involvement than the number of organizations the person belonged to, or than the tendency to vote and pay attention to political campaigns (i.e., political involvement). Also, the number of organizations joined was felt to be more indicative of social involvement than political involvement was. As a result, the scale of societal involvement was a weighted one, with social recreation having a weight of *3*, number of organizations joined a weight of *2* and political participation a weight of *1*.

Dependability. The employee's dependability and performance as rated by his or her supervisor were highly correlated ($r = .66$). Since much of the correlation between the two was probably due to the halo effect which is often present in supervisory ratings (Campbell, Dunnette, Lawler & Weick, 1970), and since the objective was to develop measures which tapped independent concepts, a derived dependability scale was formed. The revised scale was formed by "residualizing" dependability--that is, taking out the convariance between dependability and the supervisor's performance rating scale (McNemar, 1955). Thus, the revised dependability scale was a measure of an employee's dependability (as rated by his or her supervisor) *over and above* the dependability that could be attributed to the employee's productivity.

The final step in developing the criterion measures was to record all of the variables so that a high score indicated effectiveness from the perspective being used. As a result of the reduction of the number of valued outcomes to be dealt with, eleven criteria of effectiveness remained.

From an employee's perspective:

- Job satisfaction
- Depressed mood
- Overall physical health
- Work-related illness and injury
- Life satisfaction
- Self-esteem

From an employer's perspective:

- Withdrawal
- Turnover tendency
- Dependability
- Performance

From a societal perspective:

- Societal involvement

THE STRUCTURE OF EFFECTIVENESS

The initial development and validation of the Quality of Employment Indicator was based on the use of a single criterion-- Job Satisfaction. Before validating the Quality of Employment Indicator against all of the effectiveness criteria and taking into account the three different perspectives, it was important to look at the relationships among the eleven different criteria. If Job Satisfaction was found to be highly related to the other ten criteria there would have been no need to validate the Quality of Employment Indicator against the latter criteria. Its validation against Job Satisfaction would have been sufficient. However, if the criterion measures did not relate strongly to Job Satisfaction, separate validations of the Quality of Employment Indicator against these measures were clearly in order.

Table 4 shows the correlations among the eleven criterion measures. While Job Satisfaction related positively to all but one of the other criteria, it is clear from the strengths of the relationships in Table 4 that effectiveness is not unidimensional. Although some of the outcomes were highly interrelated, many were independent of each other. The implication of these results is that quality of employment measures must be validated against each criterion separately.

But, before moving to the validation of the Quality of Employment Indicator, it was useful to explore the nature of the relationships among the criteria a little further. Two questions were examined. What is the structure of the relationships among the criteria? Is the structure stable among different subsamples of people?

The first question asks whether the criteria are directly related, inversely related, or independent. If they are related positively, there is a good possibility that all criteria can be caused by the same set of working conditions. If they are negatively correlated, it could be expected that it is difficult or impossible to effect all eleven outcomes simultaneously and salubriously by improving the same working

Table 4

Pearson Product-moment Correlations Among the Criteria of Effectiveness[a]

		1	2	3	4	5	6	7	8	9	10
1.	Job Satisfaction										
2.	Depressed Mood	.45*									
3.	Overall Physical Health	.07	.09								
4.	Work-related Illness and Injury	.14*	.09	.10*							
5.	Life Satisfaction	.31*	.38*	.04	.10*						
6.	Self-esteem	.25*	.21*	-.04	.01	.29*					
7.	Employee Withdrawal	.22*	.08	.11*	.02	.06	.07				
8.	Tendency to Turnover	.46*	.25*	.03	.09	.24*	.15*	.32*			
9.	Dependability	.14*	.17*	.08	.15*	.16*	.01	.19*	.07		
10.	Performance	.21*	.12*	.02	.03	.13*	.13*	.13*	.02	.02	
11.	Societal Involvement	-.03	.10*	.09	-.01	.19*	-.08	-.12*	-.06	.04	.11*

---

*$p<.05$

[a] All criterion measures were scored so that a high numeric score indicated a "good" outcome, e.g., high satisfaction, little depressed mood, good physical health.

conditions. In order to improve some of the outcomes, others would have to be made worse. If the eleven criteria are independent, improving one outcome by altering working conditions will not affect the others.

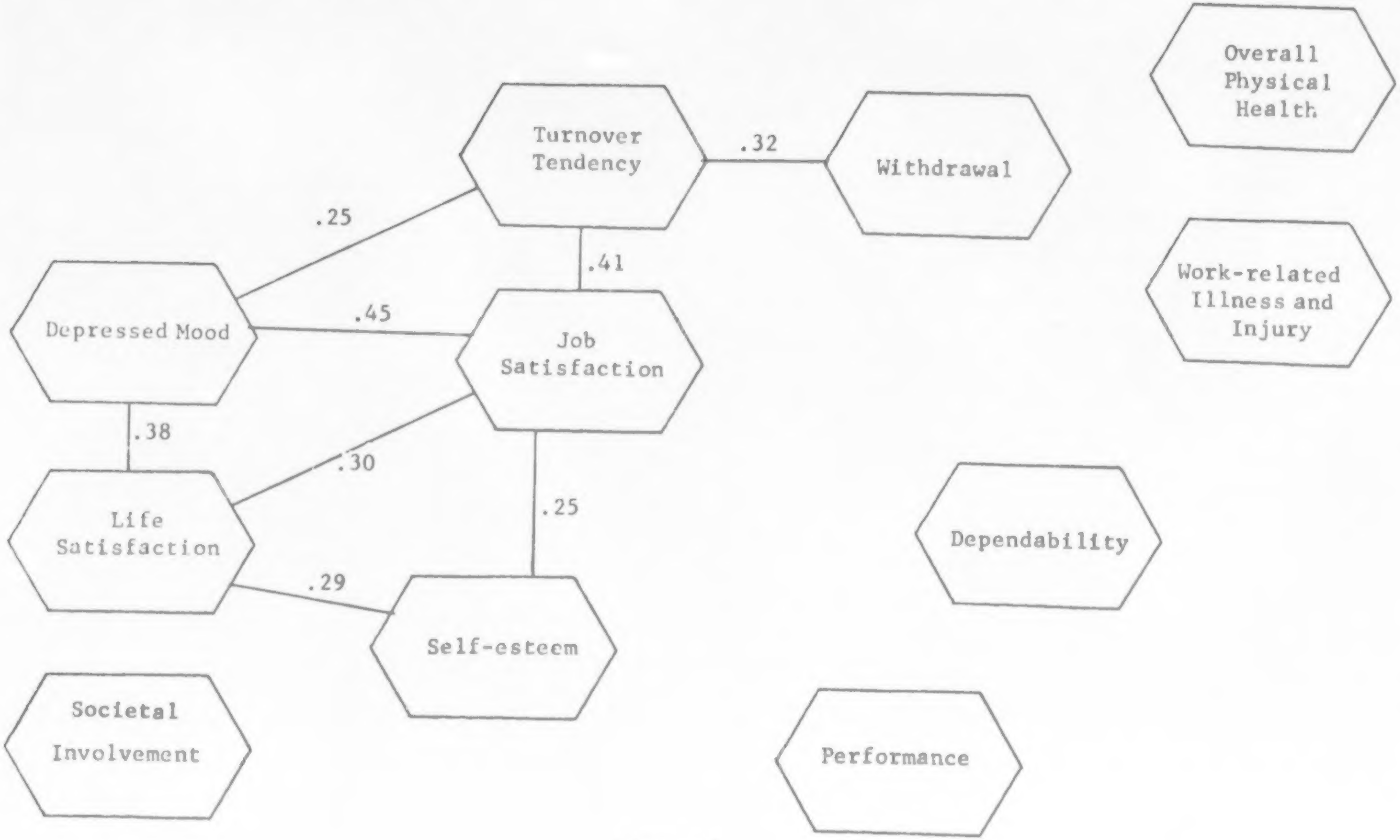
The second question is useful because the causal relationships between working conditions and outcomes may vary for different populations (e.g., men and women), or in different situations (e.g., different companies). If the structure of the relationships among the outcomes differed substantially for different subsamples, the causal relationships between working conditions and the outcomes might also have varied for the different subsamples. As a result it would have been necessary to validate the Quality of Employment Indicator for each of the subsamples.

What is the structure of the relationships among the criteria?

One means for illustrating the structure of the relationships among the criteria is a correlogram (McQuitty, 1957). The correlogram displays visually the strengths of the product-moment correlations among the variables and allows the reader to see the structure of the relationships. Figure 1, a correlogram showing the relationships among the criteria, indicates that there appeared to be a satisfaction "cluster" among the criteria. This cluster included the person's satisfaction with the job (Job Satisfaction), satisfaction with oneself (Self-esteem), satisfaction with life in general (Life Satisfaction), and feelings on the job (Depressed Mood). Turnover Tendency was also part of this cluster, and Withdrawal was in turn associated with Turnover Tendency. The fact that these two variables were in the satisfaction cluster is not surprising, since these behaviors are in part the result of dissatisfaction (Mangione, 1973; Porter & Steers, 1973).

Figure 1 also shows that societal involvement, performance, dependability, injuries and illnesses, and health are unrelated either to the satisfaction cluster or to each other.

The four psychological states embodying the employee's perspective were interrelated, but they were separate from the physiological states



Overall Physical Health
Turnover Tendency
.32
Withdrawal
.25
.41
Work-related Illness and Injury
Depressed Mood
.45
Job Satisfaction
.38
.30
Life Satisfaction
.25
Dependability
.29
Self-esteem
Societal Involvement
Performance

Figure 1

Correlogram of Relationships among Effectiveness Criteria [a]

[a] Only relationships based on r≥.25 are included

embodied in that perspective. Turnover Tendency and Withdrawal, measures of the employee's attempts to get out and stay out of the organization, were related to the employee's psychological state. These two variables are, however, measures of effectiveness from an employer's perspective. The other employer's criteria of effectiveness, Dependability and Performance, were independent of all other criteria. The societal perspective, represented by Societal Involvement, was unrelated to the other perspectives. In summary, the societal criteria of effectiveness were independent of the employee's perspective and the employer's perspective, while conceptually distinct segments of the employee's perspective and the employer's perspective were positively correlated with each other.

Is the structure of the relationships stable?

In order to examine the stability of the structure of relationships among the criteria the sample was split in two different ways, and the subsamples were examined for each method of splitting. First, the sample was split according to age. The split was made into three roughly equal subsamples: 27 years old or less (N = 193); 28-39 years old (N = 140); 40 years old or older (N = 172). The second split divided the sample into men (N = 234), married women (N = 147), and single women who had never been married (N = 79).

The correlogram in Figure 2 shows the relationships among the criteria for subsamples defined by age. Data for all three subsamples are shown on the same correlogram to highlight their similarities and differences. Figure 3 shows the relationships for the subsamples defined by the sex and marital status splits. A number of patterns emerge in these two figures. First, certain relationships appeared to be relatively stable. These are summarized by Figure 4, a correlogram wherein a line is drawn only if the relationship between the criteria was greater than .25 in at least two of the subsamples defined by either age or sex and marital status. As Figure 4 illustrates, there appeared to be a consistent set of relationships among the criteria in the satisfaction "cluster" and a

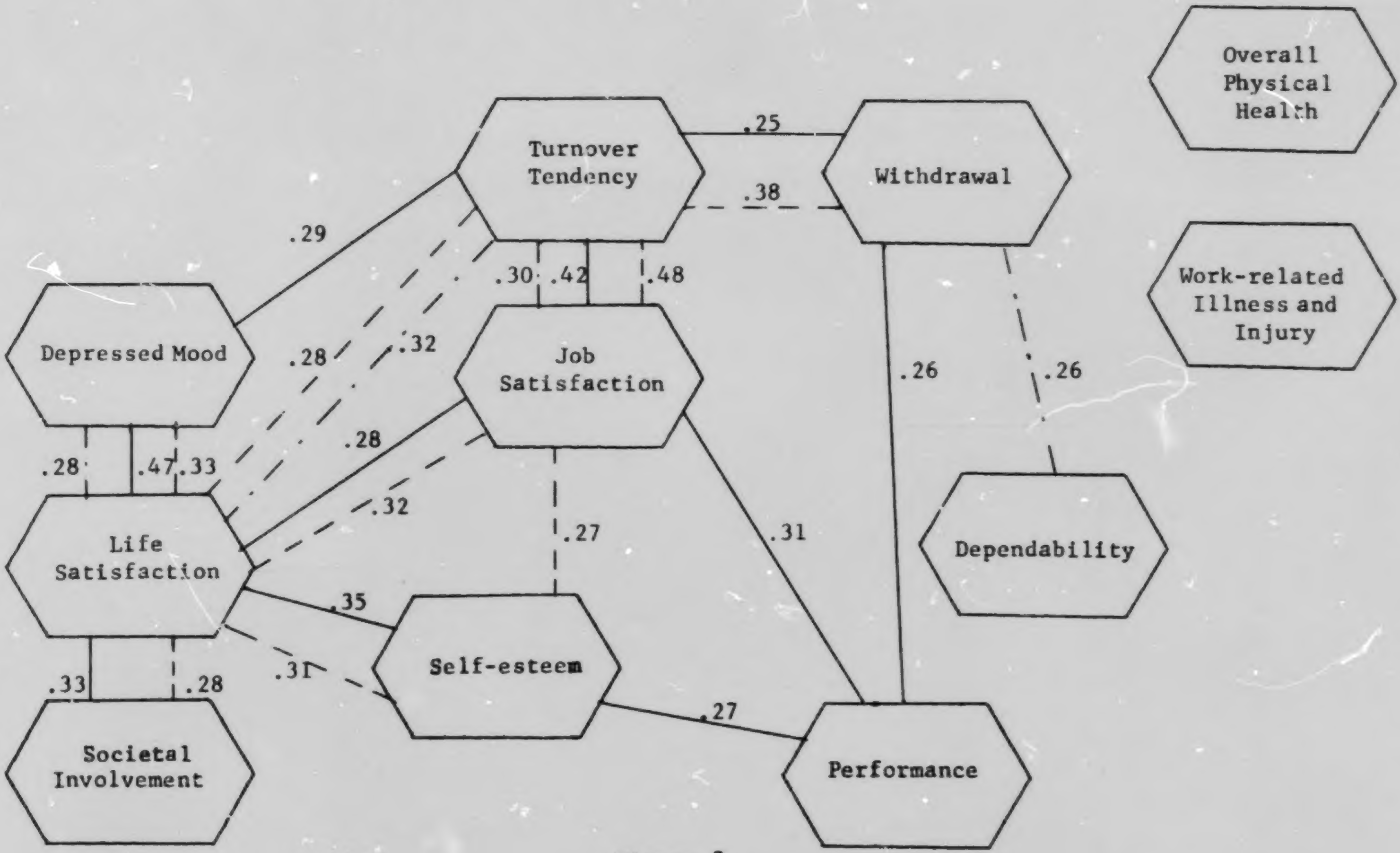

Overall Physical Health
Turnover Tendency
.25
Withdrawal
.38
Work-related Illness and Injury
.29
.30
.42
.48
Depressed Mood
.28
.32
Job Satisfaction
.26
.26
.28
.28
.47
.33
.32
Dependability
Life Satisfaction
.27
.31
.35
Self-esteem
.33
.28
.31
.27
Societal Involvement
Performance

Figure 2

Correlogram of Relationships among Effectiveness Criteria[a]

—— = young workers --- = middle aged workers -·- = older workers

[a] Only r's ≥.25 are included

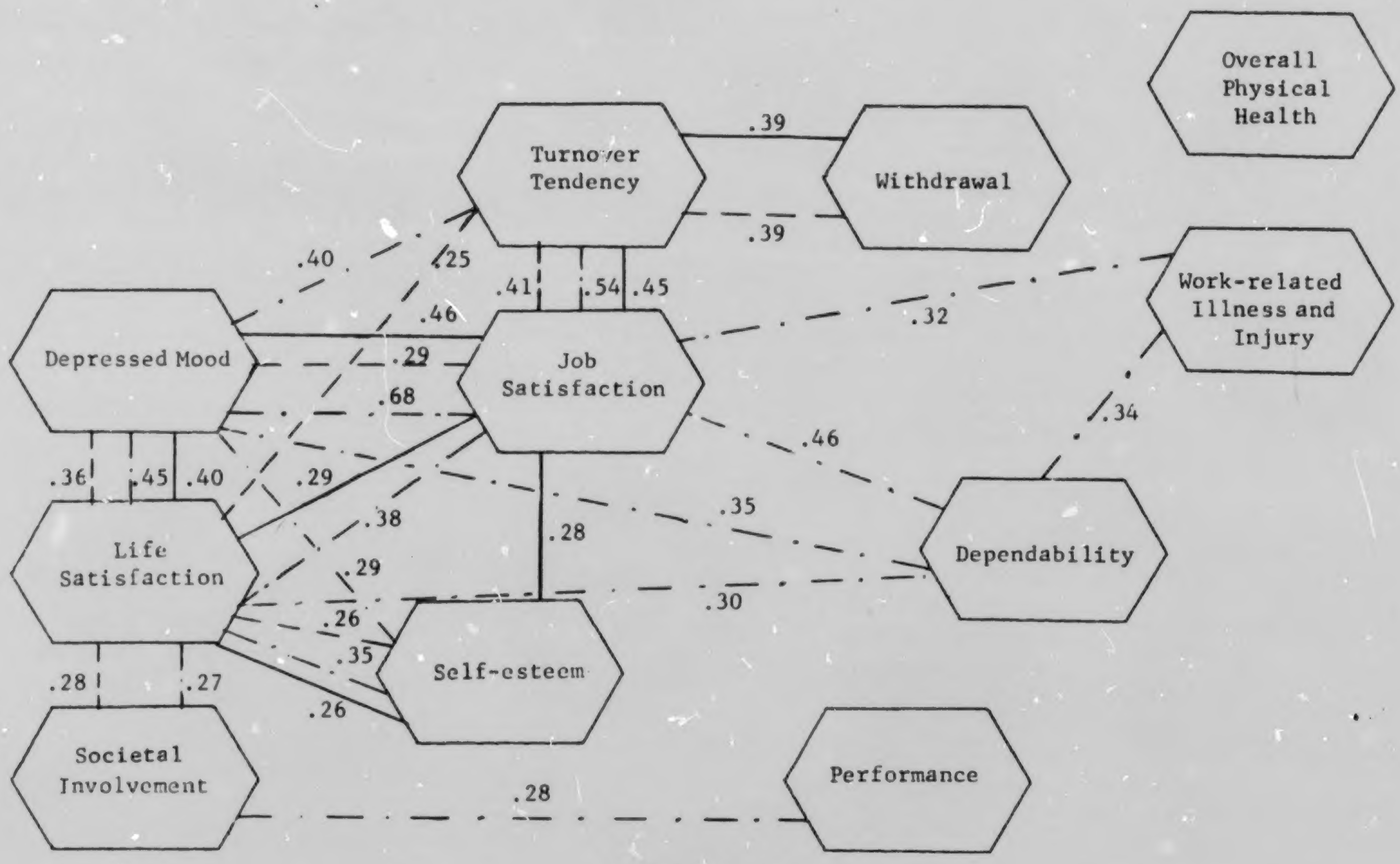

Overall Physical Health
Turnover Tendency
.39
Withdrawal
.39
.40
.25
.41
.54
.45
Work-related Illness and Injury
.32
.46
Depressed Mood
.29
Job Satisfaction
.68
.34
.46
.36
.45
.40
.29
.35
.38
Life Satisfaction
.28
Dependability
.29
.30
.26
.35
Self-esteem
.28
.27
.26
Societal Involvement
Performance
.28

Figure 3

Correlogram of Relationships among Effectiveness Criteria[a]

—— = men --- = married women -.- = single women

[a]Only r's ≥ .25 are included

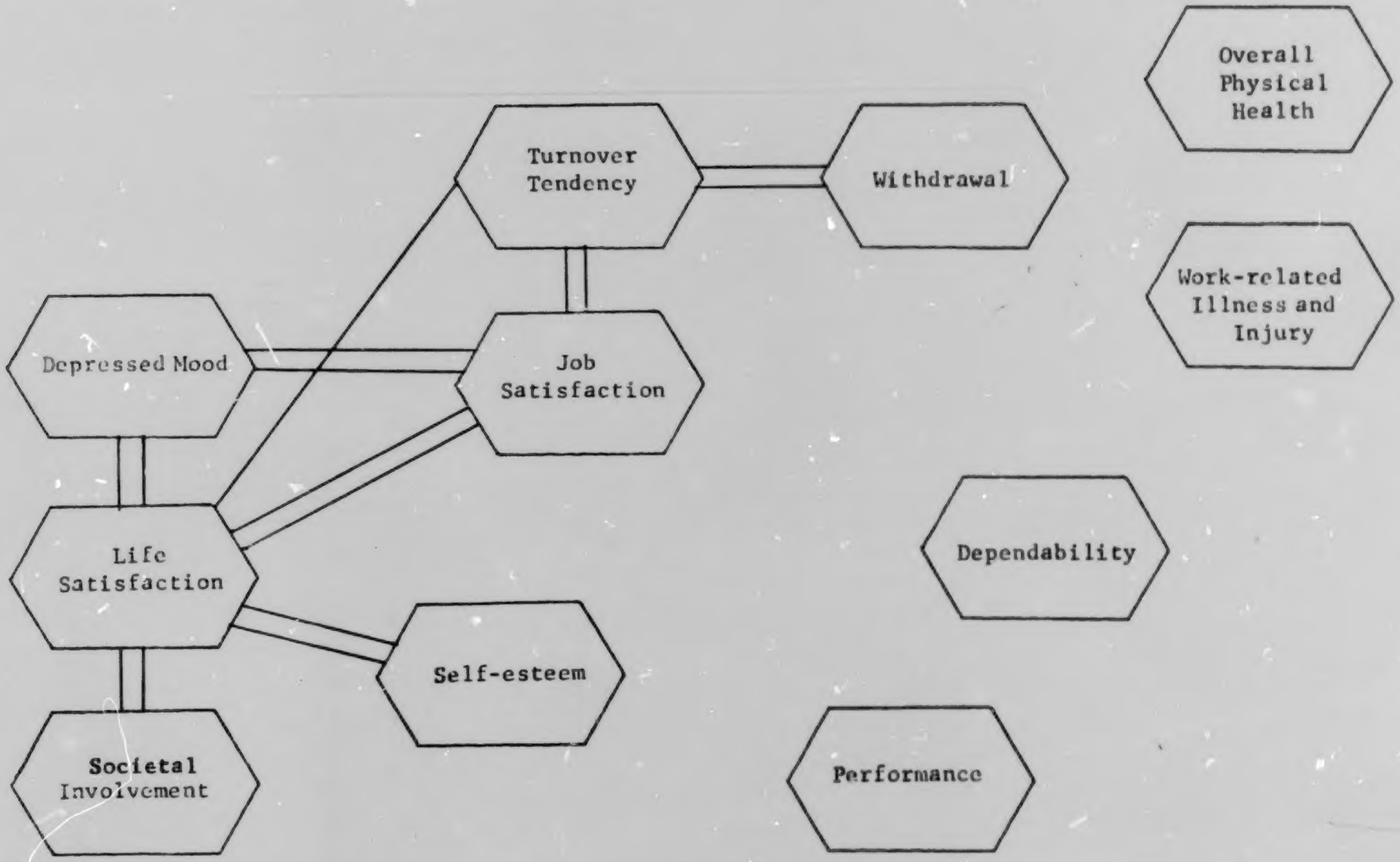

Overall Physical Health
Turnover Tendency
Withdrawal
Work-related Illness and Injury
Depressed Mood
Job Satisfaction
Life Satisfaction
Dependability
Self-esteem
Societal Involvement
Performance

Figure 4

Correlogram of Relationships among Criteria which Exist for more than one of the Subsamples Defined by either the Age or Sex

[a] Only r's ≥.25 included

fairly consistent relationship between Life Satisfaction and Societal Involvement.

But the data also showed that, in spite of these consistencies, there were considerable differences among the subsamples in the structure of the relationships among the criteria. For example, more of the criteria were interrelated among young workers and among single women than for the other subsamples. This indicates that, for these subsamples only, there may be some grounds for pooling the criterion measures into a single, overall criterion of effectiveness. However, for older workers and for men the criteria seemed much less closely related, indicating that effectiveness for these subsamples cannot be well described using one or two criteria, and multiple criteria are therefore necessary for assessing their work role effectiveness.

Overall, the pattern of relationships among the criteria indicated that all of the criteria should be used to validate the Quality of Employment Indicator. There was sufficient variation in these relationships in different subsamples to indicate that effectiveness must be viewed as both multi-dimensional and variable.

Since the Quality of Employment Indicator was specifically designed to predict job satisfaction, it was therefore inferred on the basis of these data that it would also be capable of predicting criteria shown above to be consistently related to Job Satisfaction, i.e., Life Satisfaction, Depressed Mood, and Turnover Tendency. The Quality of Employment Indicator was consequently expected to be, at best, a weak predictor of the other criteria. Whether this expectation was fulfilled will be examined later in this chapter.

Are there other criteria of effectiveness?

Thus far, it seems that effectiveness is not unidimensional and that multiple criteria of effectiveness need to be used to validate any indicators of quality of employment. At this point it is therefore appropriate to ask whether the criteria used in this study represented an adequate sampling of all possible criteria of effectiveness. The answer to this question is probably "no." The choice of the criteria used in this study was limited in a number of ways.

Limitations of sites. The criteria used were limited by the information available at the employing establishments studied. As a result, criteria could not be developed for objectively measuring individual performance, because these sites did not collect these data and limited resources made it impossible for the study staff to develop these measures.

Limitations of design. The study was designed to include a few sites and to be based on measures at one time. As a result, aggregate data (such as company profitability) and longitudinal data (such as turnover) could not be collected or used as criteria.

Limitations of conceptualization. The criteria used were based on a non-random sampling of all criteria possible. Criteria were used which seemed obvious and reasonable to the researchers. However, no attempt was made to develop a complete list of possible criteria, only one or more that represented each of the three perspectives.

These limitations most seriously affected the ability to assess and analyze effectiveness from the employer's perspective. If, for example, more sites had been employed, aggregate productivity data from the sites could have indicated which working conditions were present in the most productive, and which were present in the least productive sites. Using each site as one case cannot reliably be done with only three sites. Lack of objective performance measures also hampered the analysis of effectiveness from the employer's perspective.

CONVERGENT VALIDITY

The notion of convergent validity was introduced by Campbell and Fiske (1959) to estimate the degree of agreement between two or more independent measurements of the same concept. Evidence of convergent validity for a scale is demonstrated by a high correlation between the scale being validated and another measure of the same concept obtained through an independent source.

A primary aim of the Effectiveness in Work Roles study was to assess the extent to which the Quality of Employment Indicators obtained from self-reports were an accurate reflection of "objective" Quality of Employment. One estimate of the accuracy of the Quality of Employment Indicators was the extent to which measures from this source (i.e., the interview) agreed with measures of the same concepts from other sources, e.g., the extent to which there was agreement between the interview's Quality of Employment Challenge indicator and another indicator of Challenge obtained through on-the-job observations. The degree of convergence thus obtained reflects the accuracy of the Quality of Employment Indicators and offers proof of one kind of validity for them.

The interview's overall Quality of Employment Indicator contained four scales, each of which in turn was an aggregation of "intermediate" indices and specific items. The attempt to validate the Quality of Employment Indicator, therefore, included two main steps:

1. the validation of the overall Quality of Employment Indicator;
2. the validation of the four scales constituting the overall Indicator, i.e., Challenge, Financial Rewards, Comfort, and Resource Adequacy.

The Quality of Employment Indicator consisted of 31 items (or combinations of items) representing different aspects of an employee's work life. Its components were further grouped into four general aspects of work: Challenge, Comfort, Resource Adequacy, and Financial Rewards. In order to construct a parallel Quality of Employment index from alternative data sources, data from observations and company records were used. For some of the items included in the index, it was possible to find exact parallels in the other data sources. However, the intent was not so much to find alternative forms of every item, as it was to arrive at an adequate representation of each of the four major content areas or job facets included in the index.

The search revealed a total of 27 items that represented somewhat adequately the four interview indices of Quality of Employment. The exclusive source of items representing the Comfort, Challenge, and Resource Adequacy facets was the on-the-job observations, while employers' records were used for the Financial Rewards index. The specific constitution of the four indices will be discussed later.

After the four indices had been constructed, they were combined to form a total Quality of Employment index representing the validation criterion for the interview's overall Quality of Employment Indicator.

The interview's overall score was the mean of the scores on 31 items tapping four job facets. The number of items that constituted each aspect varied, so that the contribution of each aspect to the total index also varied. In order to represent adequately the differential contributions of the four aspects to the validation index, each of the facets was assigned a weight determined by estimating the correlations between each of the interview aspects and the total interview index. For example, the interview Challenge index correlated .84 with the total interview index, while Resource Adequacy correlated .51. Therefore, the validation Challenge index was assigned a weight of .84 and the validation Resource Adequacy index a weight of .51 in constructing the total validation index. The respective weights for the Financial Rewards and Comfort indices were .55 and .48.

The validation index thus constructed correlated .45 with the interview Quality of Employment Indicator. This is shown graphically in Figure 5. In other words, the two measures shared 20.3 percent common variance. For two indices that were designed to be alternative measures of the same phenomenon, these figures are somewhat low. One possible reason for the low correlation was the reliability of the measures. But even after adjustment for attenuation due to the internal consistency

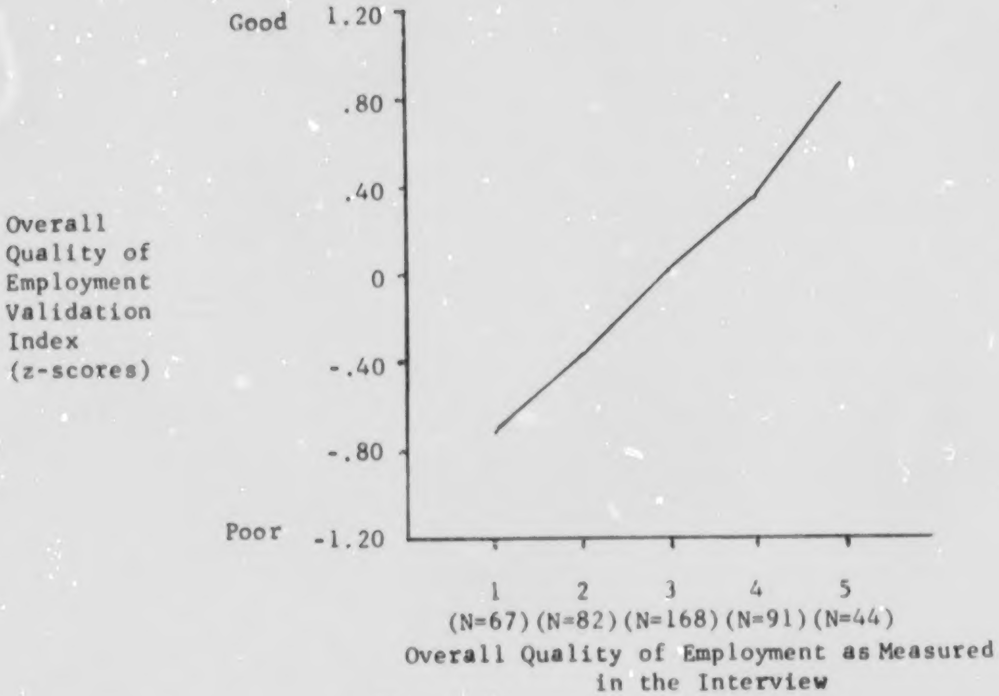


Figure 5

Association between Overall Quality of Employment as Measured in the Interview and the Validation Quality of Employment Index

Note.--In this and in later figures using the same format the following conventions were adopted.

The interview's Quality of Employment measure is presented along the abscissa and the criterion variable along the ordinate.

To facilitate comparison among criterion measures, each is presented as a *z*-score, that is, standardized with reference to its own mean and standard deviation.

The interview Quality of Employment measure was collapsed from a continuous distribution into five class intervals in a way that preserved as closely as possible the shape of the original distribution. A *1* indicates "poor" quality of employment, and a *5* indicates "good" quality of employment.

Statistical tests were based on one-way analyses of variance. The five levels of Quality of Employment as measured in the interview were the independent variable in these analyses, and the continuous criterion variables were the dependent variables. One by-product of these analyses of variance was a *eta* coefficient, a non-directional measure of association which does not require associations to be linear.

Summary Statistics: eta=.45; F=29.1; df=(4;447); $p<.001$

unreliability of the validation index, the correlation was raised only to .47. The reasons for adjusting the correlation for the reliability of the validation, but not for the interview index were two-fold. First, since the purpose of the analysis was to validate the interview's Quality of Employment Indicator as it was, not as it might have been with higher reliability, adjustment for unreliability in the interview index would have been unwarranted. Second, the validation index was, for the most part, comprised of observation indices, which were a methodological novelty, and which, therefore, warranted adjustments in anticipation of future improvements in the indices.

The adjusted correlation of .47 between the interview's Quality of Employment Indicator and the validation index indicated that the two indices still shared only 22.0 percent common variance. It remained possible, however, that some of the job facets had been better tapped than others through the relatively more "objective" data from observations and company records. If this were the case, it was probable that the moderate correlation between the two overall indices was due to the lack of agreement between measures of specific job facets rather than to the invalidity of the entire interview index.

Therefore the analysis next examined the extent to which each of the four interview measures of Challenge, Resource Adequacy, Comfort, and Financial Rewards, could be validated.

Challenge

Thirteen items from the Observation Booklet measured the degree of challenge in the job being observed. The validation Challenge index included a few items that had direct parallels in the interview and others designed to tap various Challenge dimensions of the job, some of which had been measured in the interview, and some of which had not been:

(-)1. The job requires an individual to do the same things over and over again.

(-)2. The job is so simple that virtually anybody could handle it, with little or no initial training.

(-)3. The job denies the individual any chance to use his/her personal initiative or discretion at work.

(-)4. The job is one that is highly predictable, and that rarely presents the individual with surprising or unexpected problems.

5. He/she is given enough freedom to decide how to do his/her own work.

6. The job requires a person to have a lot of skill to do it adequately.

7. The job provides an individual the opportunity to do a number of different kinds of things at work.

8. The job allows an individual to make a lot of decisions on his/her own.

9. How much variety is there in the job?

10. How much autonomy is there in the job?

11. To what extent does the job require the use of sophisticated or complex skills?

12. How much control does the employee have in setting the pace of his/her work?

13. How intellectually demanding is the job?

Negative signs indicate the items that were reversed before they were included in the Challenge scale. The first eight items were scored on a six-point scale ranging from "very untrue" to "very true." The last five were rated on seven-point scales anchored by descriptions and examples at points one, four, and seven, with a greater numeric value always indicating greater Challenge.

The correlation between the interview and the validation Challenge indices was .42; since this index was highly reliable (alpha = .98) the adjusted correlation between the two was also .42 (Table 5).

Table 5

Pearson r's between Interview and Validation Quality of Employment Measures, Together with Reliabilities of Validation Indices

Quality of Employment measures	Correlation between interview and validation indices		Internal-consistency reliability of validation index (coefficient alpha)
	Unadjusted	Adjusted for unreliability of validation index	
Challenge	.42	.42	.98
Resource Adequacy	.04[a]	.05[a]	.63
Comfort	.13	.16	.69
Financial Rewards	.37[b]	--[c]	--[c]
Overall Quality of Employment	.45	.47	.92

[a] These two correlations were not statistically significant beyond the .05 level.

[b] Excludes people working less than 35 hours a week.

[c] Construction of the Validation index did not justify an estimate of internal-consistency reliability.

Resource Adequacy

Four items from the Observation Booklet were included in the validation Resource Adequacy index.

1. He/she had adequate access to machinery, tools or other equipment.
2. He/she is given enough space to do his/her job.
3. How adequate are the resources available to the employee for him/her to do the job well?

(-)4. The work of the individual on this job was interrupted due to lack of adequate tools, information, or other resources.

The negative sign before the fourth item indicates that its values were reversed before it was included in the Resource Adequacy index. The first two and the fourth items were scored on a six-point scale. ranging from "very untrue" to "very true." The third item was rated on a seven-point scale with definitions and examples anchoring points one, four, and seven.

Table 5 indicates a low correlation between this validation Resource Adequacy index and the interview's Resource Adequacy index (.04). When the correlation was adjusted for attenuation due to lack of internal consistency in the validation index, it became .05. The variance shared by the two items was practically zero (0.3 percent).

One of the reasons for this low correlation was that the validation index consisted of only a small subset of the items in the interview index. The validation Resource Adequacy index also had low reliability of another type--low inter-rater agreement. Each of the four items that constituted the validation Resource Adequacy index showed only moderate agreement between observers who rated the jobs at the same time and almost no agreement between observers who rated the job at different times. These statistics suggest that the adequacy of resources available to an employee is not very reliably assessed through observations of the kind used in this study. An accurate estimation of Resource Adequacy appears to require more familiarity with the job than an observer was able to acquire in the one-hour observation that he or she conducted for each job. The respondent, on the other hand, knows the job well and knows what is needed to perform adequately on the job. The likelihood of the respondent being able to arrive at an accurate estimate of the adequacy of resources for the job is, therefore, high.

This argument suggests that the low convergence between the interview and the validation Resource Adequacy indices may be due in large part to the instability of its components or limitations of the on-the-job observation instrument. Until the validation instrument has been further refined and improved, it is therefore difficult to assess adequately the convergent validity of the interview's Quality of Employment Resource Adequacy index.

Comfort

The validation Comfort index consisted of eight items and a derived measure that combined the number and severity of dangerous and unhealthy conditions present in the respondent's work environment. The source for all nine measures was the Observation Booklet:

1. His/her work area is clean.
2. He/she is given adequate lighting for his/her particular job.

(-)3. His/her job exposes him/her to dangerous or unhealthy conditions.

4. He/she has enough time to do what he/she is expected to do.

(-)5. To what extent do other people make conflicting demands/requests of the employee?

(-)6. The individual doing the job is asked to do excessive amounts of work.

7. The individual working on this job is free from conflicting demands that others may make of him/her.
8. How comfortable is the physical work environment?

(-)9. The number, up to a maximum of 25, of dangerous and unhealthy conditions present.

The negative signs indicate the items for which the scores were reversed before including them in the Comfort index. The first four, the sixth, and the seventh items were scored on a six-point scale ranging from "very untrue" to "very true." The fifth and the eighth items were scored on a seven-point scale anchored by examples and definitions at points one, four, and seven.

The validation Comfort index correlated .13 with the interview's Comfort index; the adjusted correlation was .16, indicating only 2.6 percent common variance. One reason for the lack of agreement between the validation and the interview Comfort indices was that the "objective" dangerous and unhealthy conditions appeared to be not so readily observable after all--at least not by those who lack the intensive training of professional health-and-safety inspectors. This problem is discussed in detail in a special appendix at the end of this chapter. Further, the agreement between observers both simultaneously and at different times was low for most of the items that went into the validation Comfort index (Table 6).

Another reason for this poor agreement can be attributed to some lack of overlap between the items constituting the two indices. It was mentioned earlier that the objective in constructing the indices was not to replicate parallel items, but to arrive at two indices tapping the same concept. The interview's Comfort index, however, covered a variety of job characteristics that were not measurable through the other two data sources. These items included such aspects of the respondents' working life as the convenience of travel to and from work, traffic problems, problems with overtime hours, and so forth. In spite of the flexible strategy used for including items in any index, this lack of overlap between items in the two Comfort indices undoubtedly contributed to the low correlation between the two indices.

Financial Rewards

The interview's Financial Rewards index consisted of three distinct groups of variables: (a) annual wages, (b) fringe benefits, and (c) job security. Of these three, no information on job security was available from the two alternative data sources. In spite of the fact that information regarding the availability of fringe benefits was obtained from

Table 6

Kappa Estimates of Agreement for Items in the Observation Comfort Index Observed at the Same and at Different Times

Item	Kappa Same time (N = 45)	Kappa Different times (N = 448)
Clean work area	.56	.29
Adequate lighting	.16	.03
Dangerous conditions (summary rating)	.21	.23
Enough time to do work	.13	.02
Conflicting demands 1	.40	.15
Excessive amounts of work	.23	.06
Conflicting demands 2	.39	.15
Comfortable environment	.37	.26
Sum of 25 dangerous conditions	.23	.06

the company records, this variable (Fringe Benefits) was not included in the validation index. The primary reason for this was the fact that the number and kinds of fringe benefits that were available were constant within each of the sites, and even across sites there was not much variance. Addition of this variable to the validation index would, therefore, have been tantamount to adding a constant.

An examination was made, however, of the extent to which respondents tended to over- or under-report the fringe benefits that their employing organizations make available to them. These data are presented in Table 8. They indicate that, while the vast majority of the respondents were aware of the fringe benefits offered by their employers, some people did not know all of the benefits that were available to them. The two fringe benefits concerning which many people in all three companies were misinformed were the availability of stock options and free merchandise or services. Even for these benefits, however, more people were aware than were unaware of their availability. Thus, the interview measure of fringe benefits was relatively valid.

With the exclusion of two variables--Job Security (due to lack of data) and Fringe Benefits (due to lack of variance)--the only remaining validation component of the Financial Rewards index was the annual wages from one's job. These validation data were obtained from company records. The correlation between the recorded and the reported annual wages was .71, and the correlation between recorded wages and the interview's Financial Rewards index was .37. Since the validation index contained only one item, a correlation adjusted its reliability could not be calculated.

Why were the two estimates of annual wages not correlated more highly than .71? The interview asked respondents for their annual, monthly, or weekly wages, and the annual wage was calculated from the answers to that question, in combination with information on how often wages were paid. In the company records the kind of information that was available varied from site to site, so that sometimes this annual wage was a calculation from the monthly, weekly or hourly wages, and other times it was taken from the W-2 form. Estimation errors in both the interview and

Table 7

Convergent Validities of Fringe Benefits Measures

Fringe Benefit	Availability for Hospital employees					Availability for Printer employees					Availability for Auto Supplier employees				
	Company records	Interview frequencies				Company records	Interview frequencies				Company records	Interview frequencies			
		Yes	No	Don't know	No answer		Yes	No	Don't know	No answer		Yes	No	Don't know	No answer
Paid vacation	Yes	178	35	0	0	Yes	128	39	1	5	Yes	99	20	1	0
Paid sickleave	Yes	206	5	1	1	Yes	120	50	2	1	Yes	116	3	1	0
Medical insurance	Yes	185	20	8	0	Yes	168	3	2	0	Yes	114	1	5	0
Life insurance	Yes	173	22	17	1	Yes	166	2	5	0	Yes	113	0	7	0
Retirement program	Yes	202	3	8	0	Yes	167	6	0	0	Yes	112	4	4	0
Training program	Yes	150	52	11	0	Yes[a]	106	57	10	0	Yes	113	7	0	0
Profit sharing	No	40	134	36	3	Yes	172	1	0	0	Yes	112	5	3	0
Stock options	No	57	124	31	1	No	30	120	23	0	Yes[a]	37	70	13	0
Free or discounted meals	No	9	203	0	1	No	2	171	0	0	No	11	108	1	0
Free merchandise or services	Yes	166	46	0	1	No	88	85	0	0	No	28	90	2	0
Day care center	No	13	167	33	0	No	0	173	0	0	No	0	119	1	0
Use of company truck or car	No	6	205	2	0	No	0	172	1	0	No	6	114	0	0

[a] Available to some, but not all employees.

Table 8

Total Annual Wages (Frequencies) Estimated from Interviews and Company Records

(N = 144)[a]

Source	Income Categories[b]	Company Records				
		$1- $4,947	$4,948- $6,903	$6,904- $9,205	$9,206 $16,109	$16,110- $97,999
Interview	$1- $4,947	2	12	1	0	0
	$4,948- $6,903	3	88	6	1	0
	$6,904- $9,205	7	29	83	13	0
	$9,206 $16,109	8	7	33	119	5
	$16,110- $97,999	2	0	1	1	23

---

[a] Respondents working less than 35 hours per week are excluded.

[b] These categories match those used in the 1969-70 Survey of Working Conditions (with adjustments for inflation), since these defined the class intervals used in the Financial Rewards index.

the validation indices, therefore, contributed to lowering the correlation between the two wage estimates.

These errors were reduced when wage estimates from each data source were collapsed into five categories. Table 8 shows the number of respondents in each wage category and for each source. While the correlation between wages estimates obtained from the two sources was high, it remained possible that people were consistently reporting wages that were higher or lower than those in company records. The correlation would be high so long as the bias was consistent among all respondents. In Table 8, the frequencies on the diagonal are large, indicating that data from the two sources were usually in agreement. The frequencies below and left of the diagonal show a slight tendency for those who misrepresent their wages to over-represent rather than under-represent them.

Consensual validation

Workers' reports of working conditions were validated in the above pages by comparing them with descriptions of the same conditions obtained from two different sources: on-the-job observations and company records. When the reports of a particular worker are considered, there remains yet a third criterion for assessing the convergent validity of his or her reports; the descriptions of working conditions provided by others who are in the same job. To the extent that workers who share presumably similar working conditions report them similarly, their reports may be considered at best to be mutually validating and at worst to reflect only shared misperceptions of these working conditions.

To assess the "consensual validity" of the Quality of Employment Indicator and its four components an estimate was obtained for each measure of the amount of agreement among those in similar jobs. This assessment was made by first dividing the sample into work-groups, where each work-group was defined as all those reporting to a common supervisor. Within each work-group the predominant occupation was then identified, using the three-digit occupation codes from the 1960 Census. Those people who were not in the predominant occupation of their work-groups were eliminated from the analysis designed to estimate consensual validity. The analysis was therefore confined to 196 employees in 69 work groups. In each work-group members of the analysis sample had identical occupation codes and reported to the same supervisor.

The consensual validity of the overall Quality of Employment Indicator and its four components was estimated by eta coefficients of association obtained in one-way analyses of variance where the independent variable was the identification numbers of the 69 work-groups thus identified and the dependent variables were the quality of employment measures. Each eta coefficient was adjusted for its unreliability due to the small number of people in each work-group and the large number of work-groups.

The eta coefficients estimating consensual validation of the quality of employment measures are shown in Table 9, together with the other estimates of convergent validity based on on-the-job observations and company records (from Table 5).

Convergent validity: Summary

The Quality of Employment Indicator showed moderate convergent validity. According to Table 9, its consensual validity was .67, and it correlated .47 with an independent measure of quality of employment estimated by on-the-job observations and employers' records.

With the exception of Challenge, the convergent validities of the Quality of Employment Indicator's four components were lower than those of the total indicator. Challenge was the component for which the convergent validities were the highest--.68 with regard to consensual validity and .42 when the validating criteria came from sources other than interviews. The lowest convergent validities were those of the Comfort and Resource Adequacy measures when the validating criteria were based upon on-the-job observations. One of the major sources of these low validity estimates appeared to be the instability of the on-the-job observations of Comfort and Resource Adequacy. The consensual validity estimates of these two job aspects were considerably higher than were the estimates employing observational criteria.

Table 9

Estimates of Convergent Validities of Quality of Employment Indices

Quality of Employment measure	Estimate of convergent validity	
	Consensual validity	Validity with respect to on-the-job observations and company records (from Table 5)
Challenge	.68	.42
Resource Adequacy	.49	.05
Comfort	.30	.16
Financial Rewards	.46	.37
Overall Quality of Employment	.67	.47

CONSTRUCT VALIDITY

The last matter to be examined in validating the Quality of Employment Indicator was its construct validity. The critical issue in determining the construct validity of any quality of employment indicator is that it must be associated with effectiveness as viewed from some perspective. The purpose in developing a quality of employment indicator in the first place was to determine the "goodness" of the working conditions being measured. Working conditions that represent high "quality of employment" must, by definition, produce positively valued outcomes as determined from some perspective. If a quality of employment indicator is not strongly associated with outcomes that are positively valued from the perspective being used, it is not a valid indication of quality of employment as viewed from that perspective.

In assessing the construct validity of this study's Quality of Employment Indicator two separate questions were asked. First, since the Quality of Employment Indicator was originally validated using job satisfaction as a criterion, could estimates of the indicator's validity with respect to that criterion be replicated? Second, could the measure be validated against each of the other criteria of effectiveness described earlier in this report?

Validation with respect to Job Satisfaction

The first test of the validity of the Quality of Employment index was to examine its association with Job Satisfaction. Figure 6 shows that this association was very strong. Although the data are not reported here, the association was also strong within all of the subsamples examined (i.e., among women, men, supervisors, non-supervisors, young, and old). These data indicate that the Quality of Employment Indicator is a valid indicator of quality of employment when Job Satisfaction is the criterion.

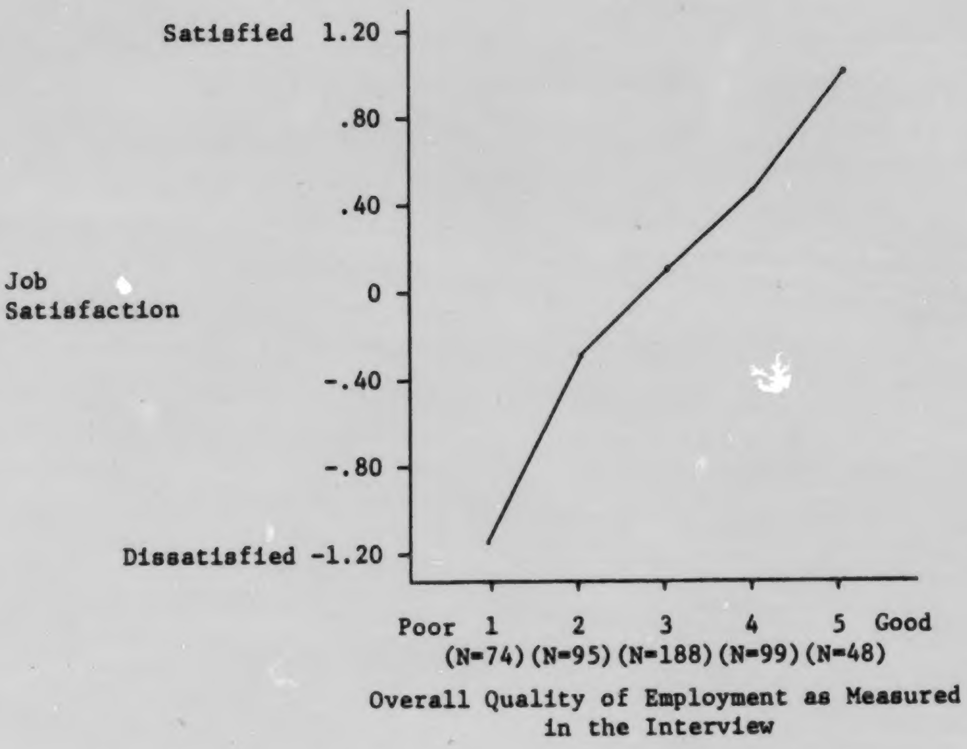


Figure 6

Association between Overall Quality of Employment as Measured in the Interview and Job Satisfaction

Summary statistics: eta=.59; F=65.13; df=(4, 499); $p < .001$

Validation with respect to Depressed Mood, Life Satisfaction, and Self-esteem

The data describing the relationships between the Quality of Employment indicator and the three affective criteria of Depressed Mood, Life Satisfaction, and Self-esteem are shown in Figures 7, 8, and 9, respectively. These data indicate that none of these relationships was as strong as that between Quality of Employment and Job Satisfaction. The relationship for Depressed Mood was nevertheless moderately strong (eta= .39) and indicates that Quality of Employment is a fairly valid indicator with regard to this criterion. Depressed Mood is probably influenced more strongly by temporary, specific job conditions, and certainly by personality factors, than is Job Satisfaction. As a result, it would be expected to be less closely related to Quality of Employment, which does not tap these factors.

Figures 8 and 9 indicate substantially lower relationships between Quality of Employment and both Self-esteem and Life Satisfaction. This was not surprising for a number of reasons. Life satisfaction is caused by a great many factors other than one's job. As a result, it relationship to quality of employment can be expected to be relatively low. The data indicated not only that the relationship was indeed low, but also that it was not monotonic. One plausible interpretation of this is that while very good and very bad quality of employment significantly influence life satisfaction, for moderate levels of quality of employment non-job factors have the major influence on life satisfaction.

There was a fairly consistent relationship between Quality of Employment and Self-esteem. The meaning of this relationship (eta = .23) is, however, not completely clear. On one hand, good quality of employment may produce high self-esteem. On the other hand, employees who have high self-esteem may appear more confident, aggressive, and competent than others. This in turn may make it more likely that they will be given challenging, high paying jobs, i.e., better quality of employment (Campbell, et al., 1970). In addition, employees with high self-esteem may prefer challenging jobs and be more likely to redefine their jobs in ways that make them challenging.

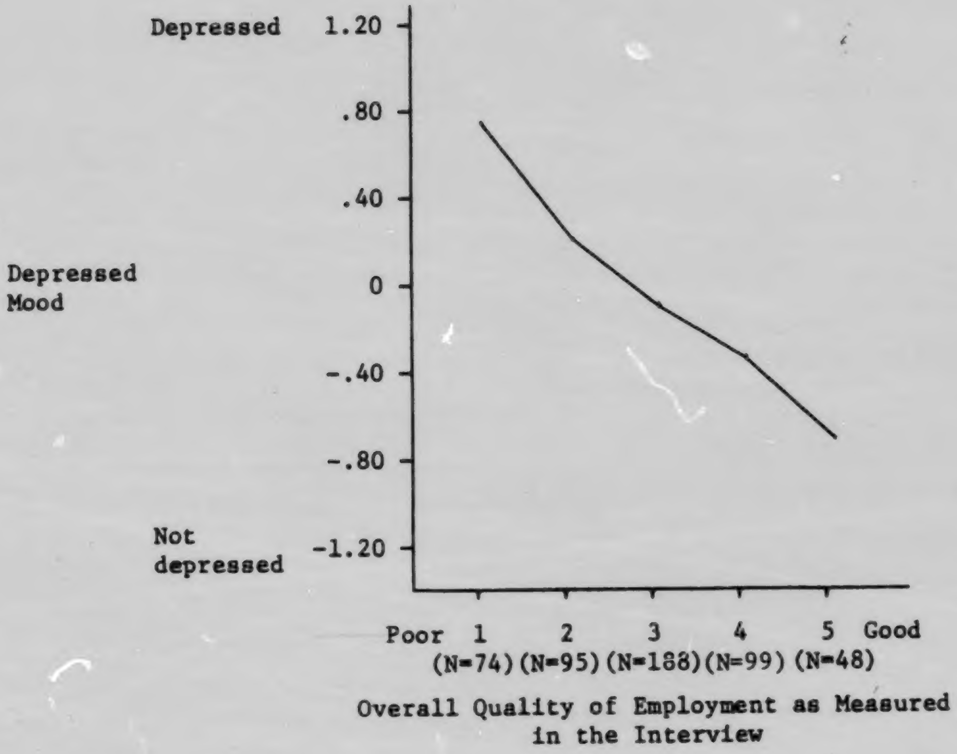


Figure 7

Association between Overall Quality of Employment as Measured in the Interview and Depressed Mood

Summary statistics: eta=.39; F=23.10; df=(4, 499); $p < .001$

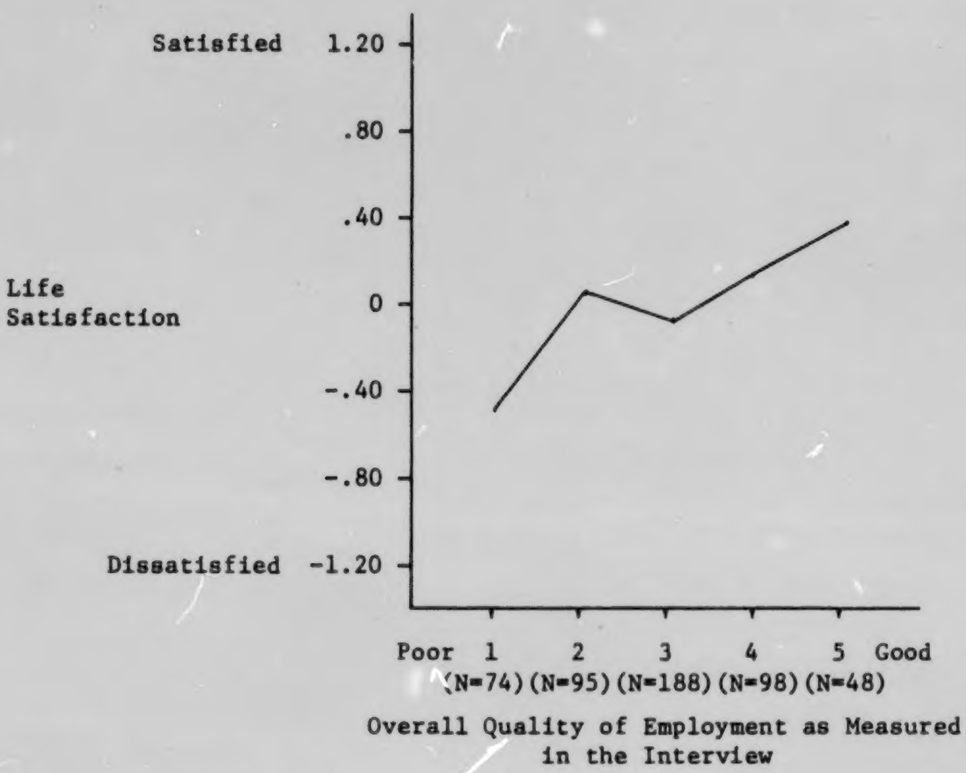


Figure 8

Association between Overall Quality of Employment as Measured in the Interview and Life Satisfaction

Summary statistics: eta=.24; F=7.66; df=(4, 498); $p < .001$

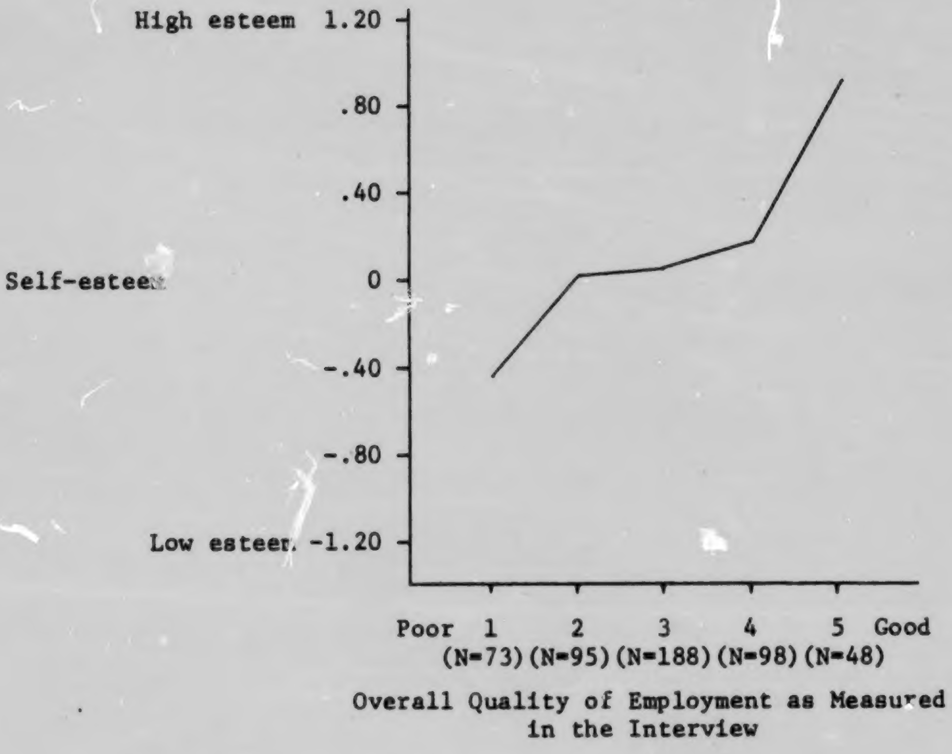


Figure 9

Association between Overall Quality of Employment as Measured in the Interview and the Validation Self-esteem Index

Summary statistics: eta=.23; F=7.19; df=(4, 497); p < .001

Validation with respect to Overall Physical Health and Work-related Illness and Injury

The relationships between Quality of Employment and measures of both Overall Physical Health and Work-related Illness and Injury are shown in Figures 10 and 11. These figures indicate that there is some validity in using the Quality of Employment indicator as a measure of good working conditions according to these criteria, but that far better measures could be developed. Jobs with very good or bad quality of employment also seemed to be appropriately "good" or "bad" according to the two criterion measures. The strengths of these relationships were not strong, however, and at intermediate levels of quality of employment, the relationships disappeared completely.

In evaluating the value of the Quality of Employment indicator for predicting Overall Physical Health and Work-related Illness and Injury, it is important to keep in mind that both of these criteria can be influenced by non-job-related factors. As a result, no quality of employment measure can be expected to relate strongly to these criteria. The data showed that while quality of employment is compatible with physical health and the absence of work-related illness and injury, better quality of employment predictors can certainly be developed.

Validation with respect to Withdrawal, Turnover, and Dependability

From the employer's perspective, working conditions will be "better" if they produce low levels of employee withdrawal (e.g., absenteeism and lateness), little tendency for employees to turnover, and high levels of employee dependability. The relationships between the Quality of Employment Indicator and these criteria are shown in Figures 12 through 14. They were all statistically significant and in the predicted direction. They were not very strong, however, and there was better predictability for extreme conditions of Quality of Employment than for moderate levels.

Considerable research (for a summary see Quinn, Staines, and McCullough, 1973) indicates that employee behaviors such as turnover and withdrawal are the result of employee's choices based on their

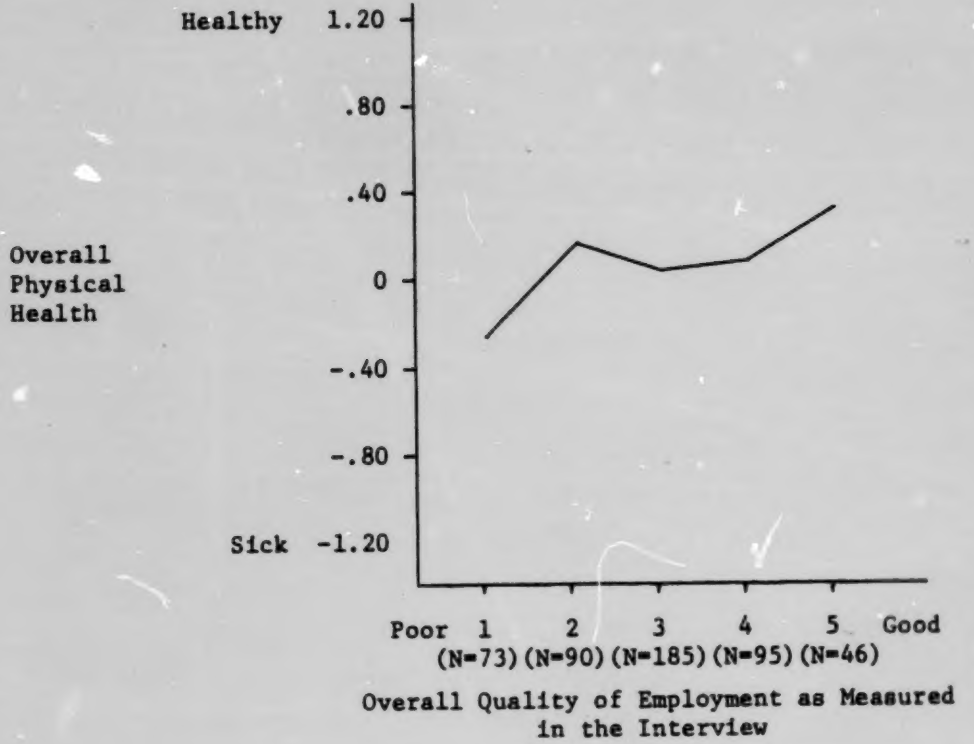


Figure 10

Association between Overall Quality of Employment as Measured in the Interview and Overall Physical Health

Summary statistics: eta=.15; F=2.93; df=(4, 498); $p < .05$

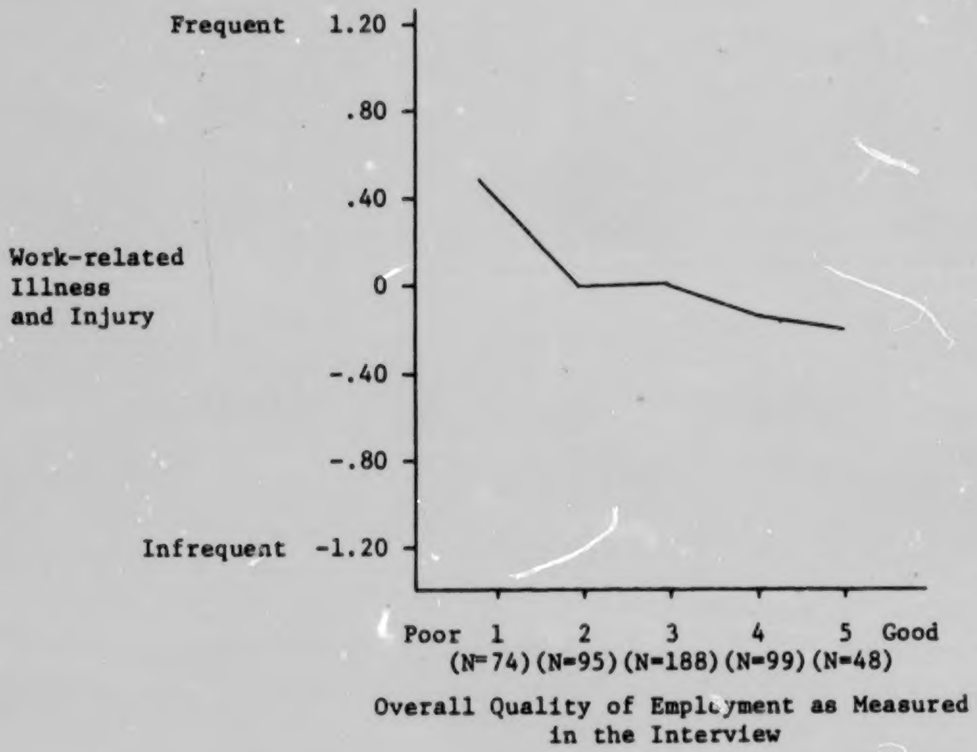


Figure 11

Association between Overall Quality of Employment as Measured in the Interview and Work-related Illness and Injury

Summary statistics: eta=.20; F=5.00; df=(4, 499); $p < .001$

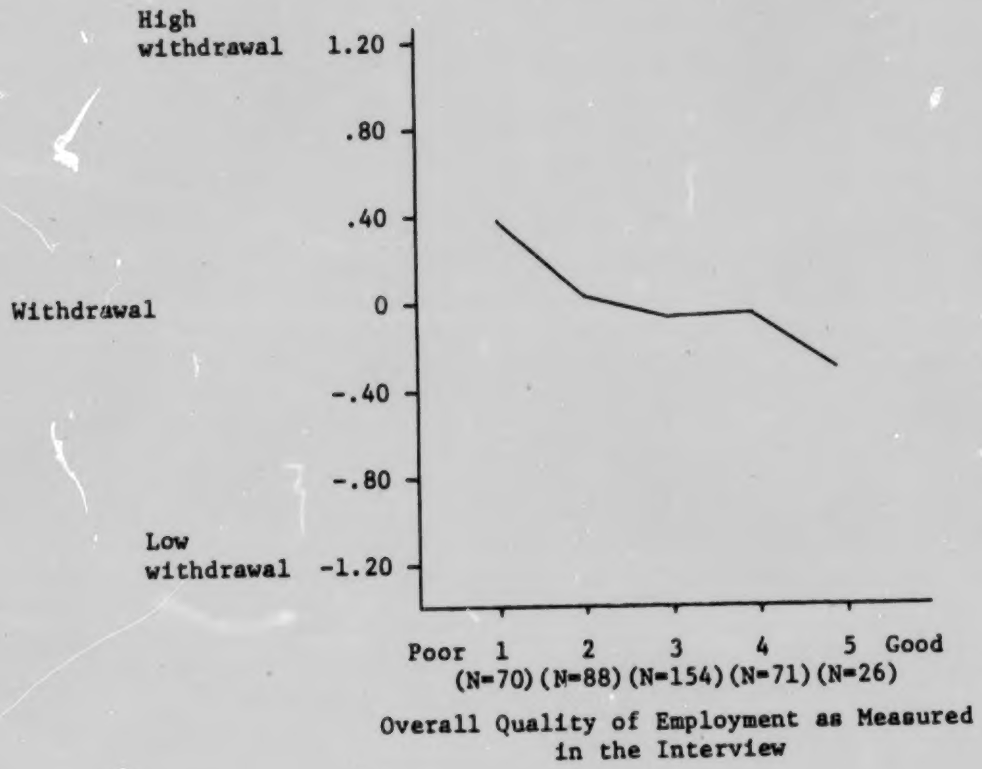


Figure 12

Association between Overall Quality of Employment as Measured in the Interview and Withdrawal

Summary statistics: eta=.20; F=4.07; df=(4, 404); $p < .01$

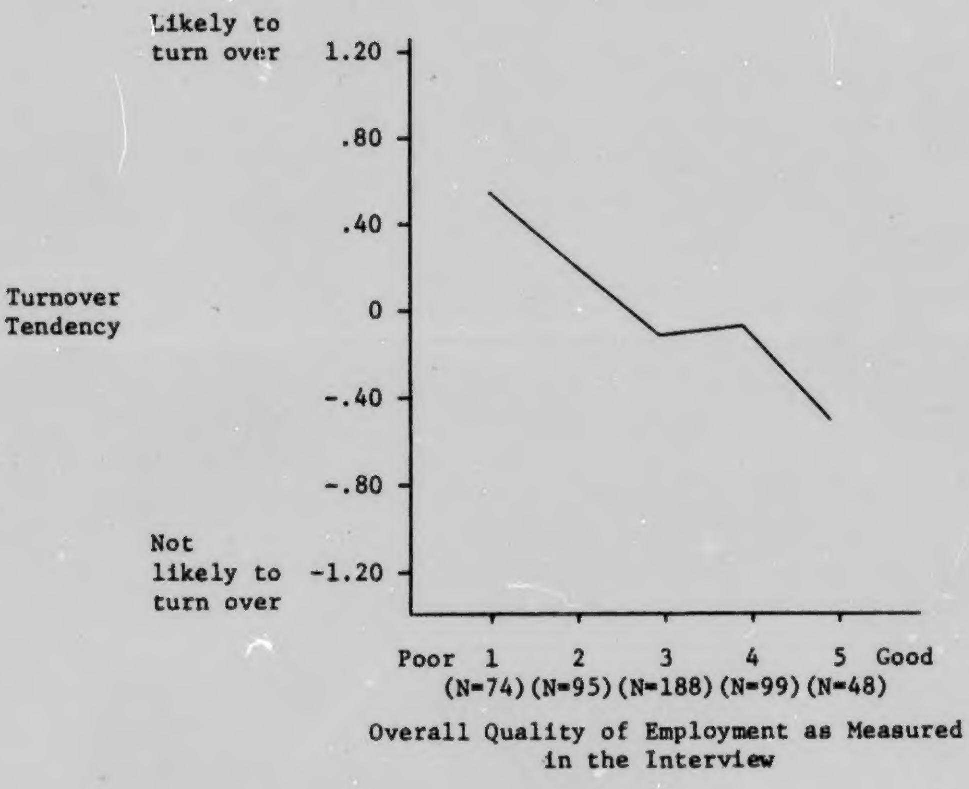


Figure 13

Association between Overall Quality of Employment as Measured in the Interview and Turnover Tendency

Summary statistics: eta=.30; F=12.28; df=(4, 499); $p < .001$

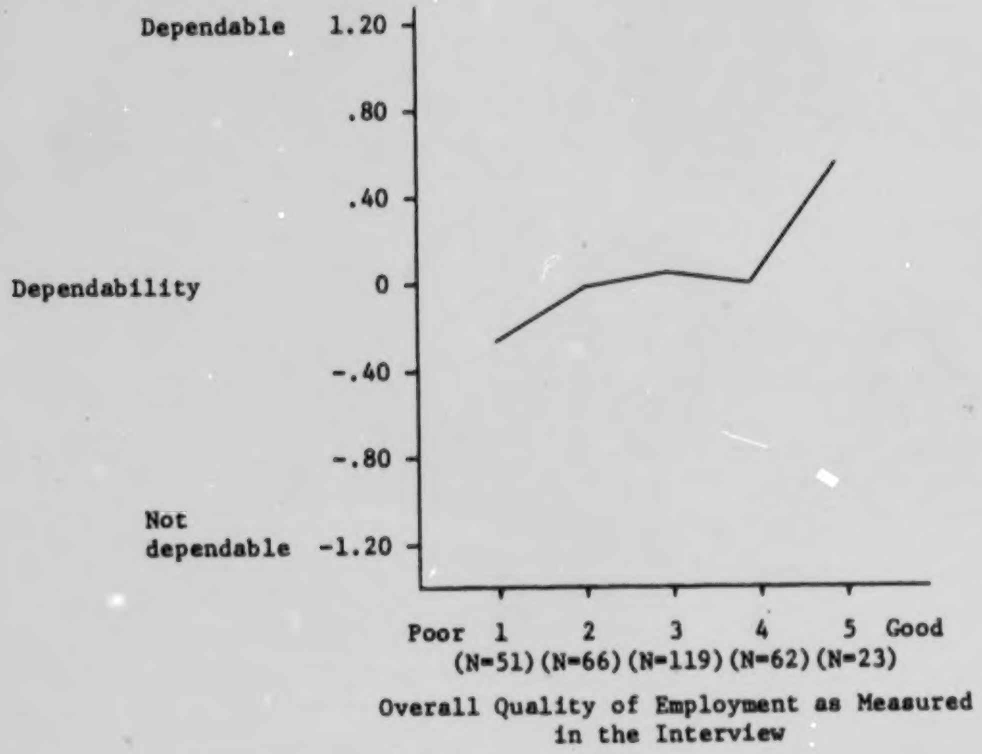


Figure 14

Association between Overall Quality of Employment as Measured in the Interview and Dependability

Summary statistics: eta=.19; F=3.04; df=(4, 316); $p < .05$

satisfaction with their jobs. According to one interpretation of these findings, if employees meet their needs and goals at work, they will not only be satisfied with their jobs, but they will also be more likely to come to work and less likely to quit. If this interpretation of the forces producing absenteeism and turnover is correct, the relationship between absenteeism or turnover and any quality of employment indicator will be limited, since job satisfaction will act as an intervening variable. This model and its implications can be seen in Figure 15. If the model shown in Figure 15 is correct, and _a_ is the relationship between Quality of Employment and satisfaction, and _b_ is the relationship between satisfaction and the criterion, then _c_, the relationship between Quality of Employment and the criterion, cannot be greater than _a·b_ (Blalock, 1964). Using a correlation coefficient as the measure of relationship and Turnover Tendency as the criterion, _a·b_ was .29 and _c_ was also .29 (note that Turnover Tendency is reverse coded so a high score means low turnover tendency), indicating that the only way to increase the relationship between Quality of Employment and Turnover Tendency may be to increase the relationship between Job Satisfaction and Quality of Employment, or the relationship between Job Satisfaction and Turnover Tendency. When Withdrawal was taken as the criterion, _a·b_ was .14 while _c_ was .20. These two correlations did not differ significantly, again indicating that better prediction of the criterion using the Quality of Employment Indicator may not be possible without developing measures of the quality of employment that are closely related to satisfaction.

The situation was less clear with respect to Dependability. While Dependability as rated by a supervisor is likely to be influenced by working conditions, it may also reflect the personalities of both the supervisor and the employee. As a result, the possible relationship between working conditions and Dependability may be limited to some extent. It nevertheless seems possible that a better Quality of Employment Indicator predicting Dependability can be developed, since the relationship between the Quality of Employment Indicator and Dependability was not high (eta = .19).

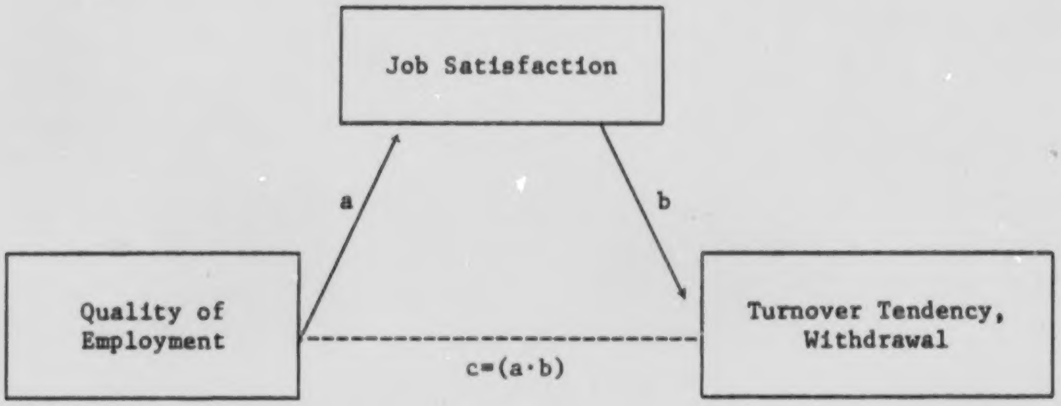

Job Satisfaction
a
b
Quality of Employment
Turnover Tendency, Withdrawal
c=(a·b)

Figure 15

Possible Relationships among Quality of Employment, Job Satisfaction, and Turnover, and Withdrawal

Validation with respect to Performance

The relationship of the Quality of Employment Indicator and the Performance measure is shown in Figure 16. Although there was a significant relationship between Quality of Employment and Performance (eta = .20), the relationship was neither linear or monotonic. Overall, the Quality of Employment index did not appear to be a particularly valid measure when it came to predicting Performance. Certainly, better quality of employment measures can be developed for predicting this criterion. Particularly important in this regard is the requirement that a quality of employment indicator cover matters more closely related to workers' motivation. Among the many characteristics of jobs or their occupants that are likely to be associated with performance, few were represented in the Quality of Employment Indicator examined in this study. Such slighted characteristics should certainly be given greater weight in future attempts to define and measure quality of employment as viewed from an employer's perspective.

Validation with respect to Societal Involvement

The relationship between the Quality of Employment index and employees' Societal Involvement is shown in Figure 17. There was a moderate, monotonically increasing relationship (eta = .22) between the Quality of Employment measure and the criterion of Societal Involvement. The data therefore indicated that the Quality of Employment Indicator was a somewhat valid measure of good working conditions when this criterion was considered. This observed relationship is particularly interesting because the societal involvement of an individual is surely the result of a great many factors other than the conditions at his or her place of work . The strength of the observed relationship was low enough to suggest that more valid measures of working conditions can be developed for predicting societal involvement. But given the nature of the criterion, the existing Quality of Employment indicator probably predicts Societal Involvement as well as any purely job-related measure is likely to do.

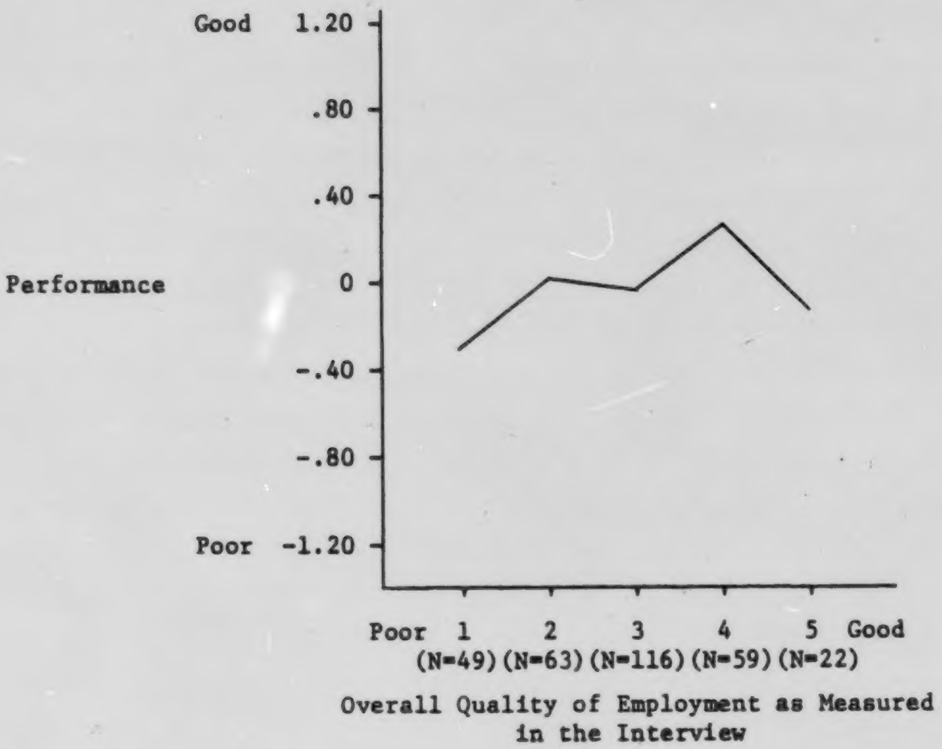


Figure 16

Association between Overall Quality of Employment as Measured in the Interview and the Validation Performance Index

Summary statistics: eta=.20; F=3.01; df=(4, 304); $p < .05$

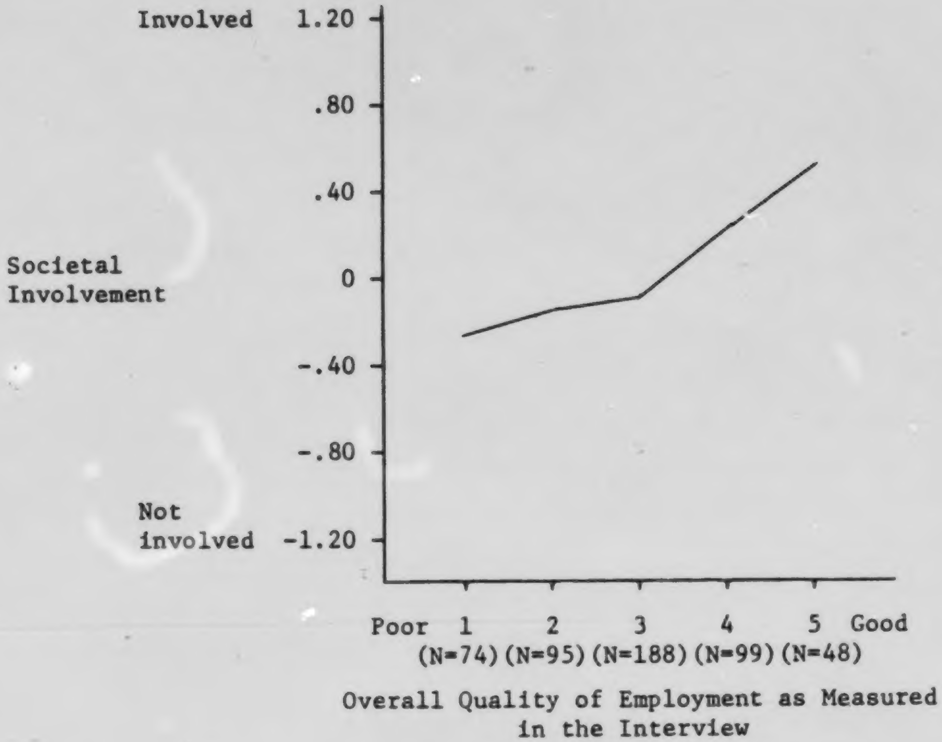


Figure 17

Association between Overall Quality of Employment as Measured in the Interview and the Validation Societal Involvement Index

Summary statistics: eta=.23; F=6.70; df=(4, 499); $p < .001$

Summary of the construct validity of the Quality of Employment Indicator

The associations between the Quality of Employment Indicator and each of the eleven criteria of effectiveness in work roles are presented in Table 10. Generally, the indicator was found to be best doing what it was originally designed to do--to predict overall job satisfaction. It predicted all other criteria less well. There were two major reasons for this diminished predictive power. First, many criteria could reasonably be assumed to be determined to a great, but unknown, extent by circumstances irrelevant to working conditions. The best illustrations of such criteria were this study's measures of physical and mental health. Secondly, even the study's job--relevant criteria were more plausibly predictable from job characteristics not included in the Quality of Employment indicator. For example, the indicator did not include measures of most of the working conditions summarized by Quinn, Staines, and McCullough (1973) as likely sources of job performance, absenteeism, and turnover.

In order to test the stability of the relationships reported in Table 10, each of the observed correlations between Quality of Employment and a criterion of effectiveness was routinely re-computed for more homogeneous subsamples of workers. These subsamples were defined according to their supervisory status, their marital status, and their age and sex. Generally the correlations observed among the full sample were successfully replicated for all subsamples thus defined. The major exception was Depressed Mood. While Quality of Employment was strongly related to Depressed Mood among younger workers and women who had never been married, the two measures were less strongly related among the other subsamples.

Table 10

Magnitude of Associations between Overall Quality of Employment as Measured in the Interview and Eleven Criteria of Effectiveness in Work Roles

Criterion of effectiveness	Magnitude of association
Job Satisfaction	Strong
Depressed Mood	Moderate-strong
Life Satisfaction	Moderate; best for extreme values
Self-esteem	Moderate; best for extreme values
Overall Physical Health	Low; best for extreme values
Work-related Illness and Injury	Low; best for extreme values
Withdrawal	Low; best for extreme values
Turnover Tendency	Moderate-strong
Dependability	Low; best for extreme values
Performance	Low
Societal Involvement	Moderate

IMPROVING THE QUALITY OF EMPLOYMENT INDICATOR

The Quality of Employment index originally had been developed to predict only one criterion of effectiveness--job satisfaction (Barnowe, Mangione, & Quinn, 1972). Data from the present study showed that the measure continued to be a good predictor of job satisfaction. Moreover, it was able, albeit less successfully, to predict other criteria of work role effectiveness.

Can the Quality of Employment index be improved? Two strategies of improvement are possible. The first involves essentially methodological improvements carried on within the conceptual language of the original measure. The "vocabulary" of such improvements would remain confined to the facets of Comfort, Challenge, Resource Adequacy, and Financial Rewards. But more powerful and efficient combinations of these facets would be explored, as well as better means of measuring them. A second strategy involves expanding the conceptual framework of the original measure. Under this strategy, information on the four "basic" facets would be supplemented with information about other job facets not included among the basic ones. These "supplementary" job facets would, moreover, be selected so as to increase the index's construct validity with regard to criteria of effectiveness other than job satisfaction. These supplementary job characteristics could include, for example, role stress (as a predictor of health) or characteristics affecting job-related motivation (as a predictor of performance).

Several analyses were, therefore, conducted in order to answer the question of whether--and to what extent--the original Quality of Employment Indicator could be improved. All of these were governed by the first of the two strategies just described. That is, the conceptual limits were never enlarged to encompass job facets other than those of Comfort, Challenge, Resource Adequacy, and Financial Rewards. Within these limits refinements of the indicator were confined to securing improved predictors of effectiveness through better combinations of or more extensive measures of the four basic facets. Three such refinements were attempted:

1. obtaining improved weightings of job facets measured in the interview;
2. supplementing interview measures of these job facets by measures of the same facets obtained by other methods;
3. reducing the number of job facets to a "most-efficient" subset.

Each of these three refinements is discussed below.

Obtaining improved weightings of job facets measured in the interview

While the Quality of Employment Indicator contained four distinct components--Comfort, Challenge, Resource Adequacy, and Financial Rewards--the four were not weighted equally in the construction of the summary measure. The different weights assigned to the four facets were not, however, arbitrary. They resulted instead from selecting only those measures of quality of employment that were appreciably associated with job satisfaction. As a result, the overall indicator included twelve measures of Challenge, nine of Comfort, five of Resource Adequacy, and five of Financial Rewards (see Table 1 above).

For predicting criteria of effectiveness other than job satisfaction, there may, however, be more powerful differential weightings of the four job facets. Indeed, even the differential weighting of the facets in predicting job satisfaction requires replication and refinement. For this reason, 22 estimates of the association between quality of employment and effectiveness were obtained. Half of these were simple correlations between the eleven criteria of effectiveness and the interview's Quality of Employment Indicator as originally constructed. These correlations (eta coefficients) are shown in the first column of numbers in Table 11. The next column, headed "R," shows the multiple correlation between each criterion and an optimum weighting of the four job facets. This optimum weighting was determined by eleven multiple regressions, one for each of the eleven criteria of effectiveness. In each of these, the four interview measures of Comfort, Challenge, Resource Adequacy, and Financial Rewards were used as predictors.

Table 11

Associations (Etas and R's) between Effectiveness Criteria and Quality of Employment (N = 505)

Criterion	Eta[a]	R
Job Satisfaction	.59	.70
Depressed Mood	.39	.41
Life Satisfaction	.24	.24
Self-esteem	.23	.23
Overall Physical Health	.15	.13
Work-related Illness and Injury	.20	.23
Withdrawal	.20	.26
Turnover Tendency	.30	.37
Dependability	.19	.18
Performance	.20	.18
Societal Involvement	.23	.23

[a] Since eta is a non-directional measure of association, all values are positive. As Figures 6-14, 16, and 17 have shown, all observed relationships were in the predicted directions.

Whenever the R was substantially greater than the eta it could be concluded that the original Quality of Employment index might be improved as a predictor of effectiveness by altering the weights assigned to its four component facets. This occurred with regard to three criteria of effectiveness: Job Satisfaction, Withdrawal, and Turnover Tendency. In the cases of the remaining eight criteria the R did not differ substantially from the eta, indicating that the original weighting of the four facets of Quality of Employment could not be improved.

Supplementing interview measures of job facets by measures of the same facets obtained by other methods

As the data in Table 5 indicates, the interview's measures of the facets of Comfort, Challenge, Resource Adequacy, and Financial Rewards were imperfect measures of these four concepts. These measures were far from perfectly correlated with alternative measures of the same things based on other methods: measures of Comfort, Challenge, and Resource Adequacy obtained by on-the-job observations, and a measure of Financial Rewards obtained from employers' records. Nor were the four latter "observational" measures themselves perfect. Since both the interview and observational measures suffered from their peculiar limitations, the question remained of whether supplementing the interview measures with observational[1] information could increase the formers' power to predict the eleven criteria of effectiveness.

To answer this question eleven multiple regressions were undertaken, one for each of the eleven criteria of effectiveness. In each there were eight predictors: the four interview measures of Comfort, Challenge, Resource Adequacy, and Financial Rewards; and the corresponding four observational measures. Since the estimate of concern was the extent to which the addition of the observational measures improved upon the predictive efficacy of the interview measures, a constraint was put on the multiple regressions. This constraint involved performing stepwise multiple regressions that gave the interview measures predictive priority. The results of these regressions, expressed as multiple correlations (R's), are shown in Table 12. The first column of numbers contains the multiple correlation between each criterion and the four interview measures without any consideration of the observational measures. Of greater interest is the second column of numbers, which shows the comparable multiple correlations obtained when interview measures were supplemented by observational measures. In order to assess the stability of the latter estimates, the initial multiple correlations were based on a randomly selected half-sample of respondents. The remaining half-sample was used to estimate the stability of the regression equations thus obtained. That is, for each criterion a set of optimal weights was obtained for the eight Quality of Employment measures, using the first random half-sample. These weights were then used to construct a regression equation, and the predicted values thus obtained were correlated with the actual values observed in the second half-sample. While the correlations in the "Unreplicated" column of Table 12 capitalize on chance relations between quality of employment measures and effectiveness, entries in the "Replicated" column do not do so.

---

[1] The term "observational" will be used in the remainder of this report to refer to any measures not obtained in the interview, in spite of the fact that the Financial Rewards measure was based upon employers' records rather than on-the-job observations.

The critical comparison in Table 12 is that between entries in the "Interview only" column and the "Replicated" column. Any indication that the latter value was substantially greater than the former means that adding observational information to the interview's measures increased the ability of quality of employment indicators to predict effectiveness in work roles. This increase in prediction was observed most conspicuously with regard to the effectiveness criterion of Societal Involvement (where an increase from .13 to .33 was observed). Smaller increases were observed with regard to Withdrawal (.10 increase) and Overall Physical Health (.08 increase). For the remaining eight criteria of effectiveness there was no advantage in supplementing interview information with observational information. In three of these

Table 12

Multiple Correlations (R's) between Effectiveness Criteria and Quality of Employment as Measured by (1) the Interview Only and (2) the Interview and Observations

Criterion	Source of Quality of Employment Measure		
	Interview only (N=253)[a]	Interview and observations	
		Unreplicated (N=253)[a]	Replicated (N=252)[b]
Job Satisfaction	.70	.71	.66
Depressed Mood	.41	.41	.37
Life Satisfaction	.25	.27	.27
Self-esteem	.26	.31	.23
Overall Physical Health	.18	.26	.26
Work-related Illness and Injury	.28	.32	.05
Withdrawal	.27	.33	.37
Turnover Tendency	.39	.44	.24
Dependability	.26	.28	.05
Performance	.28	.28	.15
Societal Involvement	.13	.37	.33

[a] Based on first random half-sample of workers.

[b] Based on second random half-sample of workers.

eight instances (Work-related Illness and Injury, Dependability, and Performance), combining interview and observational data provided predictors of effectiveness that as a set were less able to predict effectiveness than was the interview information by itself.

Reducing the number of job facets to a "most efficient" subset

How efficient is the Quality of Employment Indicator, where "efficiency" is defined as obtaining the greatest amount of predictive power at the least cost? In other words, were some of the quality of employment facets pulling more than an equal share of predictive weight, perhaps even rendering some of the other facets superfluous--and hence inefficient?

To answer this question, one step-wise multiple regression was obtained for each criterion of effectiveness. The predictors were the eight job facets measured by both interviews and observations. No constraints were put on the sequence in which the predictors entered the regression, so that each of the eight measures of job facets had an equal chance of being selected as the "single best predictor" of a criterion. At each step of the regression the value of the R obtained was compared with that obtained in the immediately preceding step, where one less predictor had been used. If the value of R did not increase between the two steps, the predictors used in the prior step were identified as the "most efficient predictors." With regard to all eleven criteria a set of "most efficient predictors" could be identified that encompassed less than the full set of predictors (see last two columns of Table 13). In no case were more than three facets included in the "most efficient" set, and in some instances only one was included. With the exception of Performance and Societal Involvement all these predictors came exclusively from the interview. In its ability to predict effectiveness with regard to a large number of criteria the most efficient predictor was clearly Challenge, with Comfort and Resource Adequacy lagging considerably behind. Financial Rewards was a poor fourth, being implicated only in Overall Physical Health and Life Satisfaction.

Table 13

Summary of Relationships between Quality of Employment and Eleven Criteria of Effectiveness

Criterion	Degree of association between Initial Quality of Employment index and criterion[a]	Could prediction be improved by assigning new weights to the four interview measures?[b]	Could prediction be improved by adding additional information from on-the-job observations or employers' records?[c]	Could number of facets be reduced without jeopardizing the prediction?[d]	Quality of Employment facets that were the "most efficient predictors" of criterion	
					Facet	Data source
Job Satisfaction	Strong	Yes	No	Yes	Resource Adequacy Challenge Comfort	Interview Interview Interview
Depressed Mood	Moderate-strong	No	No	Yes	Resource Adequacy Challenge Comfort	Interview Interview Interview
Life Satisfaction	Moderate	No	No	No	Resource Adequacy Financial Rewards	Interview Interview
Self-esteem	Moderate	No	No	Yes	Challenge	Interview
Overall Physical Health	Low	No	Yes	No	Financial Rewards	Interview
Work-related Illness and Injury	Low	No	No	Yes	Challenge Comfort	Interview Interview
Withdrawal	Low	Yes	Yes	No	Challenge Comfort	Interview Interview
Turnover Tendency	Moderate-strong	Yes	No	Yes	Challenge	Interview
Dependability	Low	No	No	Yes	Challenge	Interview
Performance	Low	No	No	No	Challenge Resource Adequacy	Observations Observations
Societal Involvement	Moderate	No	Yes	No	Challenge Comfort	Observations Observations

[a] For details see Figures 6-14, 16, and 17.

[b] For details see Table 11.

[c] For details see Table 12.

[d] For details see Table 14.

However, the selection of these "most efficient predictors" at times amounted to selecting only the most efficient of a poor lot. Table 14 compares the R's based on eight predictors with the unreplicated R's based on the "most efficient predictors" as listed in Table 13. The multiple regressions producing the latter R's provided the information for selecting the "most efficient predictors." These regressions were based on a random half-sample (Table 14). In some instances the regression weights thus obtained failed to replicate on a second random half-sample. The "most efficient predictor" columns of Table 13 should therefore be read in conjunction with the data in the column headed, "Could the number of facets be reduced without jeopardizing the prediction?" Only where a "yes" is entered in this column does the identification of "most efficient predictors" provide reliable information, since the replicability of other sets of "most efficient predictors" was low.

Focusing only on the most reliable tests of efficiency makes the pervasive importance of Challenge even more conspicuous. Challenge emerged as a "most efficient predictor" in all six instances where the number of predictors could be reduced without jeopardizing prediction. In three of these instances (i.e., with regard to Self-esteem, Turnover Tendency, and Dependability) it was the sole efficient predictor.

In conclusion, this study's attempts to improve the ability of the Quality of Employment indicator to predict effectiveness in work roles pushed the indicator as far as it reasonably could be expected to go. Better means of combining its elements were explored. Supplementary data sources--on-the-job observations and employers' records--were used, and a reduction of the measure's components to a more efficient subset was attempted. The results of these efforts were summarized in Table 13. With regard to nine of the eleven criteria of effectiveness in work roles some improvement in the Quality of Employment indicator was possible. In most cases, however, the improvement was only marginal. Those criteria of effectiveness that were only weakly associated with Quality of Employment continued to be so in spite of some evident improvement.

Earlier in this report two strategies for improving the Quality of Employment indicator were distinguished. The first involved essentially

Table 14

Multiple Correlations (R's) between Effectiveness Criteria and Quality of Employment Measured by both Interview and Observations Using (1) the Full Set of Eight Quality of Employment Measures and (2) the Reduced Set of "Most Efficient Predictors" for Each Criterion

Criterion	All Eight Quality of Employment predictors--replicated (N=252)[a]	"Most efficient" Quality of Employment predictors	
		Unreplicated (N=253)[b]	Replicated (N=252)[a]
Job Satisfaction	.66	.71	.69
Depressed Mood	.37	.39	.39
Life Satisfaction	.27	.23	.11
Self-esteem	.23	.22	.21
Overall Physical Health	.26	.16	.03
Work-related Illness and Injury	.05	.28	.17
Withdrawal	.37	.25	.26
Turnover Tendency	.24	.37	.27
Dependability	.05	.23	.10
Performance	.15	.22	.07
Societal Involvement	.33	.33	.27

[a] Based on second random half-sample of workers.

[b] Based on first random half-sample of workers.

methodological refinements instituted within the conceptual framework of the original indicator's--that is, maintaining Comfort, Challenge, Resource Adequacy, and Financial Rewards as the measure's only building blocks. The analyses reported above adopted this strategy exclusively. Its potential appears to have been exhausted. Future analyses must adopt the alternative strategy of expanding the conceptual base of the measure to encompass job facets better attuned to predicting the full range of criteria of effectiveness in work roles.

REFERENCES

Barnowe, J. T., Mangione, T. W., & Quinn, R. P. The relative importance of job facets as indicated by an empirically derived model of job satisfaction. Ann Arbor: Survey Research Center, 1972.

Belloc, N., Breslow, L., & Hochstim, J. Measurement of physical health in a general population survey. *American Journal of Epidemiology*, 1971, *93*(5), 328-329.

Blalock, H. M., Jr. *Causal Inferences in Non-experimental Research*. Chapel Hill: University of North Carolina Press, 1964.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. E. *Managerial Behavior, Performance, and Effectiveness*. New York: McGraw-Hill, 1970.

Herrick, N., & Quinn, R. The Working Conditions Survey as a source of social indicators. *Monthly Labor Review*, April 1971.

Mangione, T. Turnover--Some Psychological and Demographic Correlates. Ann Arbor: Survey Research Center, 1973

McNemar, Q. *Psychological Statistics*. 2nd ed. New York: Wiley: 1955.

McQuitty, L. L. Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies. *Educational and Psychological Measurement*, 1957, *17*, 207-220.

Porter, L. W., & Steers, R. M. Organizational, work and personal factors in employee turnover and absenteeism. *Psychologcial Bulletin*, 1973, *80*(2), 151-176.

Quinn, R. P., & Mangione, T. W. *The 1969-70 Survey of Working Conditions: Chronicles of an Unfinished Enterprise*. Ann Arbor: Survey Research Center, 1973.

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey: Descriptive Statistics, with Comparison Data from the 1969-70 Survey of Working Conditions. Ann Arbor: Survey Research Center, 1974.

Quinn, R. P., Seashore, S., Kahn, R, Mangione, T., Campbell, D., Staines, G., & McCullough, M. Survey of Working Conditions. Washington: U.S. Government Printing Office, 1971. Document 2916-0001.

Quinn, R. P., Staines, G. L., & McCullough, M. R. "Job Satisfaction: Is there a trend?" (Manpower Research Monograph No. 30, U.S. Department of Labor, Manpower Administration). Washington, D.C.: U.S. Government Printing Office, 1974.

Zung, W. R. K. A self-rating depression scale. Archives of General Psychiatry, 1965, 12, 63-70.

Appendix

Dangerous and Unhealthy Conditions

Within the context of validating the Quality of Employment Comfort Index, one subsection of the index was explored in some depth. The subsection dealt with the presence or absence of various dangerous and unhealthy conditions in the respondent's work environment. A list of these dangerous and unhealthy conditions appears in Table 1.

These twenty-five types of dangerous and unhealthy job conditions were rated for presence versus absence and for severity by both the respondent and the observers. The interview data were recoded into two major categories for each of these conditions: one indicating that the respondent said the condition did not exist on his or her job (Absent); and the other indicating that he or she said the dangerous condition existed and was either "no problem at all," "a slight problem," "a sizeable problem," or "a great problem" (Present). The observational data were recoded into the same two major categories.

Although the (five) observers' rating categories (absent; present, but no problem; present, a slight problem; present, a sizeable problem; and, present, a great problem) for the dangerous or unhealthy conditions

Table 1

Twenty-five Dangerous and Unhealthy Conditions

---

1. Inadequate protective equipment or clothing: inadequate face shields, skid-resistant shoes, body covers, safety glasses, face shields, etc.

2. Inadequately guarded electrical apparatus: ungrounded or uninsulated apparatus; uncovered connections, wires, or switches, etc.

3. Unlabeled or inadequately labeled materials or chemicals

4. Inadequate guards on machinery, equipment, or tools

5. Inadequately repaired, or defective tools, machines, or equipment

6. Although employee is not normally exposed to hazards, he/she is not adequately kept away from or warned about areas where hazardous conditions exist

7. Other hazards attributable to inadequate procedures, equipment or protection not otherwise included in 1-6 above

8. Dangers from exposure to animals. Record exposure to rats or other vermin under 17 below and not here

9. Dangers from exposure to people (co-workers, customers, patients) who could do violence or abuse. Record exposure to communicable disease under 25 below and not here

10. Any other dangers from animals or people not included in 8 or 9 above

11. Inherently dangerous materials: fire; chemicals; gases, fumes; radiation

12. Inherently hazardous equipment, tools, or machines: machines or tools that could cut, mangle, chop off fingers, etc.

13. Inherently hazardous methods or procedures: working at heights, etc.

Table 1 (continued)

Twenty-five Dangerous and Unhealthy Conditions

---

14. Working with materials which are not inherently hazardous but which could be so when present in great quantity: dust, lint, smog, etc.

15. Having to do physical tasks that exceed what appears comfortable for the employee: lifting very heavy objects; extraordinarily rapid motion

16. Inadequate human or machine help in performing physical activities, such as lifting, moving, etc.

17. Poor sanitation: dirty toilets, rats, vermin, etc.

18. Slippery floors or footing: due to disrepair, grease, oil, water, excessive waxing, torn carpeting, worn stair-treads, etc.

19. Excessive noise

20. Extremes of temperature or humidity; too hot, too cold; drafty; too dry; too stuffy

21. Inadequate space: inadequate aisle space, exits, clearance for moving objects or persons; overcrowding

22. Placement hazards: things badly piled or placed; materials inadequately insured against shifting or falling.

23. Natural hazards: terrain; exposure to the elements

24. Transportation hazards experienced while going to, or from, or around on the job

25. Exposure to communicable diseases

were the same as those used in the interviews, a category may not have had the same meaning to both observers and respondents. The reason for this was the difference in prior experiences between the two sets of people. The respondents had much exposure to their jobs but no previous exposure to the rating categories and were untrained in rating the working conditions in question. The observers' situation was the opposite. They had no previous exposure to the particular jobs rated, but they had been specially trained in using the rating categories and were presumably sensitive to the presence of dangerous or unhealthy conditions.

On the basis of the recodes, 25 two-by-two tables of the following form were constructed and the appropriate frequencies were entered into the cells. Each of these tables involved the presence or absence of a particular dangerous or unhealthy condition.

Frequency of Dangerous or Unhealthy Working Conditions

		Observation Data	
		Absent	Present
Interview Data	Absent	x	x
	Present	x	x

Dangerous and unhealthy working conditions were expected to be very infrequent in the employing establishments studied, and the 25 tables showed that this was indeed the case for the respondents interviewed. In only three of the tables did at least five percent of the respondents (25 of 506) report that the dangerous or unhealthy condition was present.

Observers, by comparison, reported the condition as present in at least five per cent of the cases for 20 of the 25 tables.

There were four dangerous conditions that no respondent reported as being present: unlabeled chemicals, inadequate warnings regarding possible work hazards, and the two kinds of dangers from animals. However, there was only one condition, the first type of animal danger, that the observers reported as completely absent.

Excessive noise as a work hazard was a particularly interesting condition (Table 2). Only eight of the respondents reported excessive noise in their work places, and the observers agreed that all eight were indeed exposed to it. However, 242 observers reported excessive noise on jobs where the respondents did not report any such noise. The combined total of 250 was the largest number of observers to report the presence of any dangerous or unhealthy working condition investigated. The most plausible explanation of this discrepancy between observers and respondents is that the workers had been exposed to the noise every day, and they eventually become so accustomed to it that they either ceased to notice it or did not regard it as a hazard. Observers, new to the work environment, were therefore more likely to notice the noise and to report it as a health hazard. This variable illustrates nicely the way in which different sources of data may give different results. It is often difficult to say which is more valid, although both viewpoints probably contribute to a knowledge and understanding of the worker's environment.

After the excessive noise, the most frequently observed hazard was inherently dangerous equipment, reported by 196 observers. Other

Table 2

Frequency of Excessive Noise

		Observation Data	
		Absent	**Present**
Interview Data	Absent	223	242
	Present	0	8

dangerous conditions observed more than 100 times were: extremes of temperature or humidity (168), inadequate space (152), slippery footing (141), inherently dangerous materials (133), placement hazards (124), and exposure to communicable diseases (108).

Because most of the dangerous or unhealthy working conditions were so infrequent, however, only those that were reported by at least ten respondents were selected for further analyses. As a result, 16 of the 25 dangerous or unhealthy working conditions were dropped: inadequate protective equipment or clothing; unlabeled chemicals; inadequate guards on machinery, equipment, or tools; inadequately repaired or defective tools; inadequate warnings about safety hazards; other hazards attributable to inadequate procedures; danger from exposure to animals, type 1; other animal dangers, type 2; working with materials which are not normally dangerous but which could be so when present in great quantity; physical tasks requiring extraordinary effort; poor sanitation; excessive noise; inadequate working space; placement hazards such as unstable piles of materials; natural hazards (terrain, exposure to the elements); transportation hazards experienced while going to or from or around on the job. For the nine hazards left, measures of agreement (kappas, discussed in Chapter 4) were computed which gave the percentage of nonchance agreement achieved (on the five-category scales) between observers in two types of situations. First, all respondents were observed on two separate occasions by different observers (Different Times). Since the observers were not simultaneously watching the respondent work, there should be some recorded difference due to actual differences in the behavior or environment, but there should be general agreement if the worker's job is fairly stable from day to day. Secondly, 45 of the respondents were observed simultaneously and independently by two observers (Same Time). Differences between those observation reports were measurement error, since the work behaviors observed were identical.

The following five of the remaining nine hazards had kappas below .30 for Same Time observations, and they were therefore eliminated from further analyses: inadequately guarded electrical apparatus; dangers from exposure to people who could do violence or abuse; inadequate human or machine help in performing physical activities, such as lifting, moving etc.; slippery floors or footing; exposure to communicable diseases. Since two observers' ratings of these hazards seldom agreed even though they were watching the same job at the same time, it was concluded that the observation technique used was unreliable as a measure of these five hazards. Of the four hazards that could be reliably observed by two observers at the same time, one, inherently dangerous methods, could not be observed with any agreement on different occasions.

After the elimination of hazards either on account of low frequency of occurrence or lack of agreement between observation measures, three were left. These are listed in Table 3 together with their kappas.

Extreme temperature or humidity was clearly the hazard with the lowest agreement among the three, while inherently hazardous materials and inherently hazardous equipment both had higher agreement. The results of validation attempts for these three observation variables are shown in Tables 4-6.

Table 4 shows the data for the inherently hazardous materials variable. The two data sources were in agreement for 72.1 percent of the jobs studied. On a more evenly distributed variable, this would have

Table 3

Kappa Agreement Estimates for Dangerous and Unhealthy Conditions

Dangerous Condition	Kappa	
	Same time	Different time
Inherently hazardous materials	.50	.24
Inherently hazardous equipment, tools, or machines	.46	.24
Extremes of temperature or humidity	.38	.15

Table 4

Frequency of Inherently Hazardous Materials

		Observation Data	
		Absent	Present
Interview Data	Absent	323	115
	Present	17	18

Table 5

Frequency of Inherently Hazardous Equipment

		Observation Data	
		Absent	Present
Interview Data	Absent	267	162
	Present	10	34

Table 6

Frequency of Extremes of Temperature

		Observation Data	
		Absent	Present
Interview Data	Absent	300	161
	Present	5	7

been a very impressive figure, but the distribution of dangerous and unhealthy working conditions required that a more detailed investigation be undertaken. While 341 cases were in agreement, an even greater figure could have been obtained if all of the observers had never reported this condition. Yet such a result would not constitute validation, since it is clear that observers would be consistently under-reporting.

It is the clustering of the majority of cases into one cell (the Absent-Absent cell) that produces a high level of agreement, and this same phenomenon calls into question the meaningfulness of that agreement. When one compares the figures across the Present row (Respondent) and down the Present column (Observer) the results are not encouraging. Thirty-five respondents said that their jobs involved working near inherently hazardous materials, but only 18 of those jobs were reported by the observers as involving such materials. Observers reported that 133 jobs required working in the presence of inherently hazardous materials, but for only 18 of those jobs did the respondents agree. Therefore, when the observer reported that inherently hazardous materials were absent, the respondent was also likely to report them absent, but, when the observer reported their presence, the respondent was not likely to report them as present.

Tables 5 and 6 show a very similar pattern for the frequencies of the other two hazards that had higher agreement between observers. For both inherently hazardous equipment and extremes of temperature the respondents and the observers agreed more on their absence than on their presence. Again, because these conditions were infrequent, the total number of agreements, although very high, was a misleading indicator of the validity of the measures, and closer inspection reveals that observers and respondents typically disagreed when observers reported the dangerous or unhealthy condition as present.

Dangerous and unhealthy conditions, while on the surface representing fairly "objective" conditions in the environment, evoked different responses for the most part from observers and respondents. There are at least three possible factors behind such discrepancies. First, as noted earlier, since respondents were accustomed to their work situations whereas observers were likely to encounter the unanticipated, observers might have been expected to perceive as dangerous and unhealthy certain conditions that respondents took for granted as unproblematic. Observers, moreover, were provided with a list of 25 dangerous and unhealthy conditions to consider. Respondents had no such list, thereby suggesting a second reason for observers to have reported more problems than respondents. On the other hand, respondents knew a lot more about their jobs than did observers and were thus likely to have been aware of some dangerous and unhealthy conditions not apparent to observers. The resultant effects of these three (and possibly other) factors could have been expected to vary across the 25 dangerous and unhealthy conditions, meaning that sometimes observers would have reported more problems, other times respondents. Yet the data show that observers consistently reported more dangerous and unhealthy conditions than did respondents. It remains for future research to determine under which conditions the observers' reports are more accurate and under which conditions the respondents' self-reports are more trustworthy.

Chapter 4

STANDARDIZED OBSERVATIONS: AN APPROACH TO MEASURING THE NATURE OF JOBS

by

G. Douglas Jenkins, Jr.
David A. Nadler
Edward E. Lawler, III
Cortlandt Cammann

ABSTRACT

In an effort to determine the usefulness of standardized job observations, 35 observers were trained to observe the characteristics of jobs. Four hundred and forty-eight employees were observed for two hours and were also interviewed. The observation measures were assessed to determine if they possessed repeatability, homogeneity, and convergence. Of the 19 job dimensions studied, 11 demonstrated repeatability and homogeneity. Six of the dimensions were tested for convergence with the interview data and four showed moderate convergence. It was concluded that job observations are a potentially useful way to measure the characteristics of jobs, but that they have significant limitations.

Chapter 4*

STANDARDIZED OBSERVATIONS: AN APPROACH TO MEASURING THE NATURE OF JOBS

Virtually all theory and research concerned with employee attitudes and behaviors assumes that they are influenced by the objective characteristics of the employee's job and work environment. While studies of organizations frequently include measures of these characteristics, generally these measures consist only of self-reports by members of the organization. This heavy reliance on self-report data has been criticized because of the problems and biases inherent in self-reports and because of the dangers of depending solely on any one measurement approach (Webb, Campbell, Schwartz, & Sechrest, 1966). It also has been noted that because they are subjective, self-report measures cannot be used in efforts to improve the quality of jobs by legislation (Lawler, 1973b). Structured observation by trained observers is a possible alternative methodology. Clearly it has some disadvantages because an observer must be present in the work place. The cost can be great--observers must be trained and supported for many hours of observation and may be a disruptive or reactive force in the work situation.

Structured observation offers, however, four potential advantages:

1. Observation allows for measurement in the actual presence of job

---

*This paper first appeared in the Journal of Applied Psychology, 1975, 60, 171-181.

behaviors and working conditions, yielding real time rather than the retrospective data often characteristic of self-reports.

2. Observation involves the direct categorizing and rating of behavior and environmental conditions by independent observers, avoiding the self-report biases which involved respondents often exhibit.

3. Observations of the same behavior or working conditions can be conducted by more than one person at the same and at different times, allowing analysis to determine the stability of ratings and the degree of agreement between different observers.

4. Observations can be structured to include category systems and rating scales that are parallel to constructs used in the gathering of self-report data, permitting the same construct to be observed by an observer and to be tapped through the self-report of an employee, thereby yielding multiple measurement data.

Two studies have used standardized job observation procedures. Hackman & Lawler (1971) observed 13 jobs in a utility company, and Turner & Lawrence (1965) observed 47 jobs in 11 companies. Both studies assumed that all individuals performing what was defined as the same job had jobs with the same characteristics. While many job characteristics are constant across the range of people who perform a given job, jobs are often tailored by management or the individual (Hackman, 1969) to fit the individual job holder. Employees who hold the same "jobs" may in fact perform very different "jobs." Thus, if objective ratings of jobs are to be maximally useful for predicting individual behaviors and attitudes, ratings of job characteristics must be made for each job holder.

These studies also leave a number of unanswered questions with respect to the usefulness of job observations.

1. Both studies involved observation of only a few job characteristics. Little is known about which job characteristics can and cannot be observed.

2. The studies presented no data on the agreement among observations made by different observers and fail to report data on the relationship between observations made at different points in time. Little is known about the stability of observational data or the levels of interrater agreement.

3. The studies presented little data on the degree to which observers and job incumbents agree on the characteristics of jobs. It is impossible to determine if the data provided by observers are likely to be simply a replication of those obtained when self-report measures are used.

4. Both studies employed organizational psychologists or sociologists. It is not clear whether useful observations can be made by individuals who are not professionals.

5. Both studies made use of observer interviews with either the job holders or their supervisors. It is not clear whether useful observations can be made when such interviews are not conducted.

Criteria for Evaluation of Standardized Observations as a Data Collection Tool

Runkel and McGrath (1972) suggest that measurement techniques should be repeatable and stable, have constituent parts which are homogeneous if that was the intent, and converge with other measures designed to measure the same phenomenon.

The repeatability and stability of a measure refer to the degree to which the same score occurs when the same object is measured twice, independently, with the same measurement procedures. In the case of standardized job observations, repeatability is the degree of agreement between ratings of two observers viewing (a) the _same job_ at the _same time_ and (b) the _same job_ at _different times_.

When job factors are measured at the same time and at two different times, there are three logical categories into which the results can fall: (a) low agreement between two observers viewing the same job at the same time and the same job at different times, (b) low agreement between two observers viewing the same job at different times but high agreement when the job is rated by both at the same time, and (c) high agreement between two observers _both_ when the ratings are made at the same time and when they are made at different times. Job factors that are not readily observable (e.g., the job attitudes of employees) should fall into the first category. Job factors that are dynamic (e.g., employee interaction, supervisory feedback) should fall in the second category. Relatively stable job factors (e.g., variety in the job) should fall into the final category.

Homogeneity reflects the degree to which there is shared variance among multiple operations that are part of the same method and that are constructed a priori to measure the same concept. Thus, homogeneity refers to the extent to which there is a concordance of results among the different operations used to measure the same concept by the same method.

Convergence also refers to the concordance of results between different measures of the same thing. It differs from homogeneity in one important way, however: Convergence is the concordance of results of multiple measures between different data collection methods. While not explicitly addressed in the definition of convergence, the obverse must also be true: There should not be a concordance of results for measures designed to measure different or unrelated concepts. Campbell and Fiske (1959) refer to this criterion as discriminant validity.

It is not expected that all job factors can be measured by a standardized job observation approach in ways that will meet the above criteria. It is important, however, to determine which job factors can be measured by this approach so that further development can proceed.

Method

Subjects

Complete data were obtained from 448 of the 506 employees who participated in the study. Subjects were from three organizations: an automobile parts manufacturer, a printing concern, and four departments of a large hospital. One of the three organizations was unionized. The sample was drawn to include a wide variety of jobs. Using Census Occupational Classifications (U.S. Bureau of the Census, 1971), most of the jobs fell into six major categories: operative (19.4%), clerical (18.8%), professional/technical (17.9%), service (16.2%), craftsmen/foremen (15.6%), and managers/officials (10.7%). Thirty-six percent of the subjects earned over $10,000 and 75% had completed high school. Ninety-one percent of the subjects were white and 62% were male.

Interviews

The first contacts the subjects had with the study were letters from the Survey Research Center, the company, and, where appropriate, their union. These letters explained the purpose of the study and asked for the subject's cooperation. After receiving the letters, subjects were contacted by interviewers and asked if they would agree to be interviewed. Over 90% agreed to participate in the study. The interviews lasted about two hours and were conducted in the subjects' homes by professional interviewers from the Survey Research Center staff. The interviewers explained to the subjects the purpose of the study and stressed that all responses would be confidential. The interviewers also explained that observations would take place and emphasized that the observer was studying working conditions, not how hard the employee worked. The subjects were then given the option of not being observed and a small number declined.

The structured interview contained 19 relevant job description items, falling into 11 a priori dimensions: variety, autonomy, task identity, task feedback, worker pace control, comfort, resource adequacy, certainty, required cooperation, external feedback, and required skills and abilities. Previous research (Hackman & Lawler, 1971; Lawler, 1973a; Vroom, 1964) has indicated that some of these dimensions have important impacts on worker attitudes and behaviors.

Selection and Training of Observers

Announcements were posted stating that observers were wanted for a research project. Applicants were told that they would have to attend a two-day training session before they were hired, but that they would be paid for attending the session. In the training session, the applicants worked on simulated observational tasks and obtained feedback on their performance. This method enabled the training activities to be used for selection purposes, since trainee performance on the simulated observation task could be used as an objective measure and predictor of observer performance on the job.

Four jobs similar to those performed by employees in the sample were videotaped prior to the training. During the 2-day training period each trainee rated the videotapes using the observation instrument. Each rating was followed by sessions where the trainees discussed their ratings. The training staff encouraged discussion where there were significant differences among the ratings of the trainees. Individuals talked about why they rated the tapes as they did, what cues they attended to, and how they reached a decision for each rating. The stated goal of these sessions was to move toward agreement among observers as each iteration of the cycle occurred.

The ratings on the final observation of tapes were collected and used as the basis for selection decisions. Analysis of scores using three different measures of deviation and agreement were performed, and permanent hiring decisions were made. Of a total of 51 trainees, 35 were hired. A refresher training session was held after observers had been in the field for three weeks. A shorter (half-day) version of the original training format was used. Based on the refresher training (and analysis of the data) two additional trainees were eliminated for marginal performance.

Observers

The observers were nonprofessionals. Almost all were college students, most of them juniors, seniors, or first year graduate students.

The group included 19 men and 16 women. They were paid $2.65 per hour and all but a few worked as observers on a part-time basis.

Observation Instrument

The observation instrument included 59 questionnaire-type items measuring a wide variety of job factors. The items are presented in Table 1 along with an indication of their a priori scale assignments. The majority of these items were responded to on 6-point Likert-type scales; the rest were responded to on 7-point anchored scales (see Hackman & Lawler, 1971).

Observation Procedure

Each respondent was observed twice for an hour. The observations were scheduled so that the two different observations were separated by at least two days, were usually done at different times of the day, and were always done by two different observers. In addition, 48 observations were done with two observers rating randomly selected employees at the same time (this was in addition to and not instead of the two observations at different times). These observations were made to determine the repeatability of the measures obtained at the same time. In all, 941 hours of observation were conducted. The observation hour was structured so that the observer spent 10 min becoming oriented to the job, 30 min observing specific job actions, and 20 min rating the job in situ. The observers then typically spent an additional 15 min away from the job completing the observation instrument.

Table 1

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Variety						.896	.963	.826
	1. How much variety is there in the job?	.670	.226	3	3			
	2. The job requires an individual to do the same things over and over again	.553	.271	3	3			
	3. The job provides an individual the opportunity to do a number of different kinds of things at work.	.634	.248	3	3			
Autonomy						.862	.960	.867
	4. How much autonomy is there in the job?	.562	.232	3	3			
	5. The job allows an individual to make a lot of decisions on his/her own.	.536	.295	3	3			
	6. The job denies the individual any chance to use his/her personal initiative or discretion at work.	.564	.231	3	3			
	7. He/she is given enough freedom to decide how to do his own work.	.448	.252	3	3			
External feedback						.514	.716	.457
	8. To what extent does the employee find out how well he/she is doing on the job from his/her supervisor or coworkers?	.206	.056	2	1			
	9. The coworkers of an individual working on the job never tell the person whether he/she is doing well or poorly.	.196	.044	2	1			
	10. Supervisors generally let a person working on the job know how well they think he/she is doing?	-.026	.097	3	1			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Task feedback						.631	.847	.656
	11. To what extent does doing the job itself provide the employee with "feedback" about how well he/she is doing?	-.026	.097	3	1			
	12. Just doing the work required by the job provides many opportunities for a person to figure out how well he/she is doing.	.124	.063	3	1			
	13. The individual can see the results of his/her work.	-.262	.125	3	1			
Rigidity						.697	.823	.697
	14. How rigid does the employee appear to be in his/her attitudes and manner of working?	.252	.075	1	1			
	15. The individual working on this job appears to be one who would have difficulty adapting to new and unusual situations.	.268	.116	1	1			
Certainty						.628 (.742)	.915 (.921)	.708 (.768)
	16. How much uncertainty is there in the job?	.500	.175	3	2			
	17. How sure does the employee seem in his/her job as to whether certain things will work as expected?	.440	.080	2	2			
	18. The job requires the individual to be prepared to handle surprising or unpredictable situations.	.456	.190	1	2			
	19. The job is one that is highly predictable, and that rarely presents the individual with surprising or unexpected problems.	.429	.235	3	3			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Certainty (continued)								
	20. The individual working on the job does tasks which are clearly defined.	.206	.168	3	1			
Conflicting demands						.782	.879	.782
	21. To what extent do other people make conflicting demands/requests of employees?	.404	.145	2	2			
	22. The individual working on this job is free from conflicting demands that others may make of him/her.	.394	.151	2	2			
Interruptions						.346	.504	.346
	23. How true is it that he/she is frequently interrupted for work-related reasons?	.512	.169	2	2			
	24. How true is it that he/she is frequently interrupted for non-work related reasons?	.520	.107	2	2			
Skills and abilities						.919	.978	.918
	25. To what extent does the job require the use of sophisticated or complex skills?	.472	.288	3	3			
	26. How intellectually demanding is the job?	.600	.307	3	3			
	27. The job requires a person to have a lot of skill to do it adequately.	.495	.289	3	3			
	28. The job is so simple that virtually anybody could handle it with little or no initial training.	.404	.293	3	3			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Worker pace control						.846 (.900)	.948 (.947)	.859 (.900)
	29. How much control does the employee have in setting the pace of his/her work?	.429	.177	3	2			
	30. The job allows the individual to determine his/her own work pace.	.426	.255	3	3			
	31. How much control does the worker himself or herself have over the pace of his/her work?	.281	.276	3	1			
Dependence						.630	.773	.630
	32. To what extent does the individual depend on his/her colleagues for doing his/her job?	.329	.084	2	2			
	33. To what degree does the employee have to depend on the work performed by someone else in order to get the materials or information he/she needs to do his/her work?	.353	.133	2	2			
Cooperation						.678	.807	.678
	34. To what extent does the job of the employee require that he/she constantly check with others and others check with him/her?	.430	.116	2	2			
	35. To what degree does the employee have to cooperate directly with other people in order to do his/her job?	.282	.088	2	1			
Work pressures						.324	.487	.324
	36. Are there any pressures for better performance over and above what is reasonable?	.168	.021	2	1			
	37. The individual doing the job is asked to do excessive amounts of work.	.226	.063	2	1			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Effort						.820	.901	.820
	38. To what extent does the employee work hard on his/her job?	.139	.130	2	1			
	39. The individual working on this job expends a lot of effort trying to perform his/her job well.	.163	.091	1	3			
Meaningfulness						$*^a$	*	*
	40. The job is meaningful.	.473	.201	1	3			
Resources						.292	.708	.315
	41. How adequate are the resources available to the employee for him/her to do the job well?	.208	.018	1	1			
	42. How true is it that he/she is given enough space to do his/her job?	.248	.147	1	1			
	43. How true is it that he/she is given adequate lighting for his/her particular job?	.164	.031	1	1			
	44. How true is it that he/she has adequate access to machinery, tools, or other equipment?	.307	.042	1	1			
	45. The individual working on his job frequently had to stop to get things that he/she needed and didn't have readily available.	.195	.049	2	1			
	46. The work of the individual on this job was interrupted due to lack of adequate tools, information or other resources.	.252	-.031	2	1			
Comfort						.260 (.686)	.671 (.815)	.363 (.686)
	47. How comfortable is the physical work environment?	.369	.262	3	3			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Comfort (continued)								
	48. How true is it that his/her work area is clean?	.561	.286	3	3			
	49. How true is it that he/she has enough time to do what he/she is expected to do?	.126	.018	1	1			
	50. How true is it that his/her job exposes him/her to dangerous or unhealthy conditions?	.210	.225	3	1			
Locus of pace control						*[b]	*	*
	51. How much control does the worker himself or herself have over the pace of his/her work?	.281	.276	3	1			
	52. How much control does his/her supervisor have over the pace of the employee's work?	.297	.088	1	1			
	53. How much control does his/her work group have over the pace of the employee's work?	.400	.058	1	2			
	54. How much control does machinery or equipment have over the pace of the employee's work?	.536	.377	3	3			
	55. How much control do customers, clients, patients, have over the pace of the employee's work?	.631	.219	2	3			
	56. How much control does the flow of work from other groups or departments have over the pace of the employee's work?	.241	.082	1	1			

Table 1 (continued)

Analysis of the Observation Items

Scale	Operation	$\kappa_{w(s)}$ $n = 45$	$\kappa_{w(d)}$ $n = 448$	Predicted Category	Observed Category	Median r	Coefficient alpha	H.R.
Task identity						.614 (.623)	.815 (.767)	.623 (.623)
	57. To what extent does the employee's most frequently performed work chunk(s) represent an "entire piece of work"?	.352	.103	3	2			
	58. The job provides an individual with the chance to finish completely any work he/she starts.	.180	.098	3	1			
	59. An individual working on the job usually can complete the entire job from beginning to end.	.456	.058	3	2			

a Single item scale. Statistics not possible to compute.

b Scale not constructed to be homogeneous. Statistics not computed.

Analytic Procedure

Repeatability. Several measures of interobserver agreement or repeatability for nominal and ordinal scales exist in the psychometric literature. The most common, percentage or proportion of agreement, suffers in that it includes agreement which could be accounted for by chance; thus can be deceptively high if the number of categories is small or if only a few categories are used with any appreciable frequency. Cohen's weighted kappa (κ_w) (Cohen, 1968; Fleiss, Cohen, & Everitt, 1969) provides a chance corrected estimate of the proportion of agreement between raters. In addition, the statistic allows for partial as well as full agreement by assigning differential penalties to deviations from perfect agreement. Given the ordinal nature of the scoring system it was decided that when there was perfect agreement among observers (i.e., Observer A rated 3 and Observer B rated 3), full agreement credit (or a weight of 1.0) would be given. When the ratings differed by only one category (i.e., Observer A scored 3 and Observer B scored 2 or 4), then one-half agreement credit would be given (or a weight of .5). Disagreements of more than one category between observers were given no agreement credit (or a weight of .0).

In general, κ_w is computed by subtracting the weighted proportion of agreement expected by chance (p_c, the sum of the weighted cross-products of the marginals) from the weighted observed proportion of agreement and dividing by the maximum chance corrected agreement possible ($1 - p_c$).[1] Theoretically, κ_w may range in value from -1 to +1, however, maximum values require identical marginals. The extent to which the marginals differ lowers the practical maximum value of κ_w (Cohen, 1960).

---

[1] For the complete formulae and computational procedures of κ_w and its variance estimates, see Fleiss, *et al*. (1969).

A κ_w value of 0 would indicate no increase in agreement above the "chance" level estimate from the cross-product of marginal proportions. Negative values of κ_w represent levels of agreement less than that to be expected by chance alone. Positive values represent levels of agreement between observers above that expected by chance.

Items on the job observation instrument were placed a priori into one of the three repeatability and stability categories mentioned earlier depending upon the characteristic being measured. They were then empirically placed using the following criteria: Category 1--κ_w (same time) $<.33$, κ_w (over time) $<.20$; Category 2--κ_w (same time) $>.33$, κ_w (over time) $<.20$; Category 3--κ_w (same time) $>.33$, κ_w (over time) $>.20$.

The κ_w levels required for different categories are arbitrary and by usual conventions may appear low.[2] It must be understood, however, that κ_w is a direct measure of the proportion of weighted agreement between observers *above* chance.[3] κ_w was set at a lower level for the over-time situation because characteristics of jobs often change from time to time.

---

[2] Any κ_w which is not significantly different from 0 at the .01 level, regardless of its magnitude, was regarded as having a value $<.20$ for categorization purposes.

[3] Care should be taken not to confuse κ_w with the Pearson r. κ_w is equivalent to the produce-moment correlation only when partial agreements are weighted in such a way that the agreement weights, w_{ij}'s are according to the formula (derived from Cohen, 1968)

$$w_{ij} = \frac{(k - 1)^2 - d^2}{(k - 1)^2} ,$$

where k is the number of possible rating categories and d is the horizontal or vertical distance from the main diagonal of the k x k matrix. The weighting scheme used in the present study will typically yield a value of κ_w which is much smaller than the corresponding value of the Pearson r.

Given the behavior sampling technique employed in this study, it is very unlikely that the behavioral demands and characteristics of a job would be exactly the same at any two random points in time.

Homogeneity

The most common measure of the homogeneity or internal consistency of a set of items which comprise a scale designed to be homogeneous is Cronbach's coefficient α (the generalized equivalent of Kuder-Richardson Formula 20). Because a major determinant of the value of this coefficient is the number of items in the scale, however, the coefficient can be very high, even in cases where the inter-correlation among the items is very low. While the scales involved in the observation instrument are relatively short (none exceed six items), it is still preferable to assess the homogeneity of the scales independent of scale length (Guilford, 1954). The Scott (1960) homogeneity ratio is such a measure. The homogeneity ratio, which ranges from 0 to 1, expresses the ratio between two terms. The first term is the difference between the actual variance of the sum of the items and the variance expected if the items were uncorrelated. The second term is the difference between the variance expected if all items were perfectly correlated and the variance expected if the items were uncorrelated. This coefficient can also be interpreted as a weighted average of the scale's interitem correlations.

Results

Repeatability

A measure is considered to exhibit repeatability if it falls into

Categories 2 or 3 in the system presented earlier. Table 1 contains the $\kappa_{\underline{w}}$ scores for observations made at the same time ($\kappa_{\underline{w}(\underline{s})}$) and at different times ($\kappa_{\underline{w}(\underline{d})}$). As can be seen, 32 of the 59 measures showed empirical agreement between observers when ratings were made at the same time (i.e., the $\kappa_{\underline{w}(\underline{s})}$ values exceed .33). Of the 12 measures that were not expected to exhibit repeatability (e.g., operations of personality characteristics and resource adequacy), 9 failed to meet the .33 standard while 3 exceeded the standard. Of the 47 measures expected to be rated consistently by observers when ratings were made simultaneously, 30 exceeded the critical $\kappa_{\underline{w}}$ of .33. The operations that failed were chiefly measures of external feedback, task feedback, work pressures, and effort.

When the ratings were made at different times, 19 of the measures showed acceptable agreement between observers. In two cases, this agreement was not predicted in advance. In the case of 12 items, this kind of agreement was predicted but not obtained. These failures primarily involved measures of feedback, worker pace control, and task identity.

Homogeneity

Table 1 also contains, when appropriate, the median interitem correlation, coefficient α and homogeneity ratio for each a priori scale. When all items of the scale are not repeatable, the indices are presented parenthetically for those scale items which are repeatable. The scales designed to tap the concepts of variety, autonomy, rigidity, certainty, conflicting demands, cooperation, required skills and abilities, worker pace control, and effort are reasonably homogeneous; that is, they have homogeneity ratios greater than .60 or 60% of perfectly homogeneous operations. For the scales of external feedback, task feedback, task identity, comfort, work pressures, interruptions, and resources, homogeneity does not exist.

Convergence

Six of the scales can be meaningfully tested for convergence since they were measured by the interview, and the observation data demonstrate both repeatability and homogeneity. Table 2 presents the multitrait-multimethod matrix for the six scales. Only those observation items that are repeatable are used to construct the scales for the matrix. Entries on the main diagonal of the matrix are coefficient α values for the individual scales within a given method. The circled correlations in the lower left of the matrix are the convergence correlations between the same constructs measured by the two methods. Four of the six constructs assessed exhibit reasonable levels of convergence: variety, skills, autonomy, and pace control. Certainty and cooperation fail to show convergence between methods. It appears that what the observers assessed as certainty is related to what the respondents viewed as variety and skills and, to some extent, autonomy in their jobs. Cooperation measured by the observation instrument did not correspond to any of the other five constructs measured by the interview.

The values in the upper left and lower right triangles (solid lines) are the monomethod triangles that contain the correlations between the constructs measured by a single method. The values in the heteromethod triangles (dashed lines) are correlations between the different constructs measured by different methods. A comparison between the data in the heteromethod triangles and the convergence correlations shows that both pace control and autonomy fail to exhibit discrimination. Both relate

Table 2

Multitrait-Multimethod Matrix (n=448)

	Interview						Observations					
	1	2	3	4	5	6	1	2	3	4	5	6
Interview												
1. Variety	(.45)											
2. Skills	.41	(.79)										
3. Certainty	-.13	.04	(.58)									
4. Autonomy	.38	.44	.20	(.68)								
5. Worker pace contrc	.27	.11	.02	.25	(*)[a]							
6. Cooperation	.21	.13	-.16	-.03	-.01	(*)[a]						
Observation												
1. Variety	.43	.43	-.12	.32	.21	.23	(.96)					
2. Skills	.42	.48	-.10	.32	.22	.23	.83	(.98)				
3. Certainty	-.39	-.40	.14	-.29	-.16	-.25	-.88	-.77	(.92)			
4. Autonomy	.39	.40	-.11	.35	.27	.22	.87	.79	-.76	(.96)		
5. Worker pace control	.37	.36	-.08	.32	.30	.18	.70	.64	-.55	.83	(.95)	
6. Cooperation	.22	.17	-.07	.12	.06	.16	.42	.30	-.42	.33	.14	(.85)

[a] Single item scale.

more highly to variety and skills as measured by the interview method than they do to the interview's attempt to assess these constructs. On the other hand, the measures of variety and skills do show acceptable levels of discrimination when the convergence correlations are compared with the correlations in the heteromethod triangles.

The correlations in the observation monomethod triangle are very high. It appears that observers have difficulty discriminating among the constructs of variety, skills, certainty, autonomy, and pace control, but do seem to be able to discriminate between those dimensions and coooperation. A comparison between the convergence correlations and the monomethod triangles reveals that the criteria for discriminant validity are met for interviews but not for observations.

Discussion

The observational technique for assessing job characteristics which was used in the present study was moderately successful. About two thirds of those operations expected to exhibit repeatability did so (Table 3 summarizes the results). With minor exceptions, scales intended to be homogeneous were, in fact, composed of homogeneous items. In four of the six instances where it was possible to test for convergence, the scales tended to exhibit convergence and some degree of discrimination.

In future investigations, improved training procedures and better operations appear to be necessary to assess job characteristics in the remaining areas. It is particularly important that better measures and perhaps better conceptualization of factors such as feedback and task identity be developed since they play important roles in many theories of

Table 3

Summary Table: The Observability of Job Dimensions

	General Repeatability of Items			
Job Dimension	At same time	At different times	Homogeneity	Convergence
Variety	Yes	Yes	High	Moderate
Autonomy	Yes	Yes	High	Moderate
External feedback	No	No	Moderate	c
Task feedback	No	No	Moderate	c
Rigidity of employee	No	No	Moderate	d
Certainty	Yes	No	Moderate	Low
Conflicting demands	Yes	No	High	d
Interruptions	Yes	No	Low	d
Required skills and abilities	Yes	Yes	High	Moderate
Worker pace control	Yes	No	High	Moderate
Required interdependence	Yes	No	Moderate	d
Required cooperation	Yes	No	Moderate	Low
Work pressures	No	No	Low	d
Employee effort	No	No	High	d
Meaningfulness	Yes	Yes	a	d
Resource adequacy	No	No	Low	c
Comfort	Mixed	Mixed	Low	c
Locus of worker pace control	Mixed	Mixed	b	b
Task identity	Yes	No	Low	c

a One item scale.

b Not constructed as homogeneous scale.

c Analysis not done because of results of earlier operations.

d Not measured in the interview.

job design (Lawler, 1973a).

The greatest difficulty with the present results is the lack of discrimination among job characteristics by the observation method. The large correlations in the observation method triangle indicate that substantial method variance was present. While the case has been made elsewhere (Hackman & Lawler, 1971) that there is an underlying ecological relationship among some of the constructs measured here, the observation instrument as constructed and as used by the observers in this study was clearly subject to a systematic source of error in the form of a halo or g-factor.

Three methodological factors probably weakened the results of this study:

1. The training was geared primarily toward obtaining interobserver agreement; consistency with conceptual constructs and previous research was of secondary importance. The current study establishes that observers can be trained in such a way that they will agree with each other and that their ratings on a priori scales will be consistent. Future efforts should build on this learning and devote time to the development of an understanding of the underlying concepts of the observation instrument (e.g., What is autonomy?, What is variety?). The objective would be to improve the ability of the observer to discriminate among constructs. This should contribute to increased convergence with other measures and to increased discriminant validity.

2. The instrument was designed to collect a large amount of data and include many exploratory and/or experimental items as well as items designed to gather data for a wide variety of other purposes. Interviews with the observers indicated that the length of the instrument led to observer

boredom and problems in keeping constructs clear. This is reflected in the patterns that emerge from the multitrait-multimethod matrix. One way of improving the quality of observations would be to shorten and simplify the instrument so that it taps the key measurable dimensions of a job in a minimum amount of time.

3. Certain jobs are undoubtedly more inherently observable along the dimensions used in this study than others. An assembly line job is probably more observable by this technique than would be the job of a financial vice president. In future investigations, analyses parallel to those of this investigation on respondents subgrouped by job type could provide useful information about the applicability of the observational technique to different types of jobs. Such analyses are not possible in the present study because of insufficient cases in many job categories. Additional data collection designed to permit such analysis by job classification is currently underway.

In addition to resolving the methodological problems, future research needs to examine how structured observations such as these relate to employees' attitudes and behavior. It would not be expected that job observations will be as closely related to performance, absenteeism and turnover as will employee attitude data since, presumably, job characteristics influence attitudes which in turn influence behavior. Certain job factors, however, should be significantly related to behavior and to attitudes concerned with satisfaction and motivation. It is crucial that these kinds of links be established. Research in organizations cannot continue to consist of many studies relating attitudes to attitudes along with a few studies relating attitudes to performance. The development of an effective observation approach appears to be one way to increase the number of studies that relate environmental and structural conditions to employees' attitudes and behavior. The importance of doing so is hard to overstate; effective interventions designed to increase organizational effectiveness require knowledge of how the objective characteristics of the work environment affect individuals.

References

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Cohen, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 1960, *20*, 37-46.

Cohen, J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 1968, *70*, 213-220.

Fleiss, J. L., Cohen, J., & Everitt, B. S. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 1969, *72*, 323-327.

Guilford, J. P. *Psychometric methods* (2nd ed.). New York: McGraw-Hill, 1954.

Hackman, J. R. Toward understanding the role of tasks in behavioral research. *Acta Psychological*, 1969, *31*, 97-128.

Hackman, J. R., & Lawler, E. E. Employee reactions to job characteristics. *Journal of Applied Psychology*, 1971, *55*, 259-286.

Lawler, E. E. *Motivation in work organizations*. Monterey, Calif.: Brooks/Cole, 1973(a).

Lawler, E. E. Quality of working life and social accounts. In M. Dierkes & R. A. Bauer (Eds.) *Corporate social accounting*. New York: Praeger, 1973(b).

Runkel, P. J., & McGrath, J. E. *Research on human behavior: A systematic guide to method*. New York: Holt, Rinehart & Winston, 1972.

Scott, W. A. Measures of test homogeneity. *Educational and Psychological Measurement*, 1960, *20*, 751-757.

Turner, A. N., & Lawrence, P. R. *Industrial jobs and the worker*. Boston: Harvard University Graduate School of Business Administration, 1965.

U. S. Bureau of the Census. *1970 census of population: Alphabetical index of industries and occupations*. Washington, D.C.: U.S. Government Printing Office, 1971

Vroom, V. H. *Work and motivation*. New York: John Wiley & Sons, Inc., 1964.

Webb, E., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally & Company, 1966.

Chapter 5

BIAS INHERENT IN OBSERVATIONAL MEASURES

by

Michael Moch
Cortlandt Cammann
Nina Gupta

ABSTRACT

This paper considers alternative explanations of two previously noted limitations of standardized observations: only moderate convergence between observational and interview measures; and lack of discrimination among observational measures of different concepts. It explores such potential explanatory factors as unreliability of measures, true ecological covariance among different job dimensions, observer bias in the form of halo effects. The analyses were performed on four scales tapping job characteristics expected to be stable. Results showed that these observation scales exhibited reasonable internal consistency, repeatability, and stability. That is, different observers at the same time agreed somewhat with each other but not enough to reject the notion of halo bias among observers; and the decrease over time in agreement among observers was not unduly large, thereby suggesting ongoing changes in job characteristics. Further analyses documented the existence of the halo bias among observers but found it not to be large. The bias did prove larger, however, for the less concrete scales. Controlling on bias, nonetheless, did not eliminate the covariance among job characteristics, indicating that jobs in the present sample were distributed along a single dimension. Thus the inability of observers to discriminate among some job characteristics suggests in part that these characteristics are highly related in the workplace.

Chapter 5

BIAS INHERENT IN OBSERVATIONAL MEASURES

An earlier paper (Chapter 4) suggested that standardized job observations provide a useful alternative to questionnaires and interviews as a source of data regarding jobs. The paper demonstrated that while alternative observational measures of the same phenomenon showed reasonable convergence, there was only moderate convergence between parallel observational and interview measures. There was, furthermore, considerable lack of discrimination among observational measures designed to tap different concepts. These findings may reflect unreliability in the alternative measures, or they may represent true ecological covariance among different job dimensions. It is also possible, however, that the findings reflect observer bias--a halo effect, or a tendency for observers to see jobs as either "good" or "bad" along all dimensions. To the extent that this is true, observational measures should be used with a great deal of care, or the effects of such bias controlled statistically.

This paper represents an attempt to explore the degree of bias or halo inherent in observational measures, as well as the extent of the ecological relationships among job characteristics. Confirmatory factor analysis (Alwin, 1974; Joreskog et al, 1970; Werts et al., 1974.) be used to explore these issues. Specifically, two areas will be treated in this paper. (1) Three scales that showed stability in the earlier study (Chapter 4) will be examined to assess their repeatability and stability in greater detail. This assessment was made previously for single items only. It is important, however, to examine the repeatability and stability of scales in addition, since observer bias, while inflating estimates of internal consistency, would lower estimates of repeatability and stability. (2) The amount of observer bias or halo in the observational measures will be assessed. Controlling for this bias, then, the relationships among different job characteristics in the sample of jobs observed will be examined.

Method

The method of data collection was described fully in Chapter 4. Only a few of the measures reported in that study were used here. The analyses, in the present case, were confined to those reliable items tapping job characteristics that the researchers expected to be stable. Four such job characteristics were, therefore, chosen for analysis: specifiability (certainty), worker pace control, skill complexity, and autonomy. While the first three scales represent specific job characteristics, the last encompasses a more global job dimension. The four scales, and the items used to tap each, are documented in Table 1.

In order to analyze the repeatability and stability of the observation scales, confirmatory factor analysis was used. This analytic approach represents a way to disentangle the effects of observer bias from those effects caused by the phenomenon in question *per se*.

Table 1

SCALE	ITEM	METHOD
1. Autonomy (AS)	How much autonomy is there in the job? (A1)	Observation
	The job allows an individual to make a lot of decisions on his or her own? (A2)	Observation
	(-)The job denies the individual any chance to use his/her personal initiative or discretion at work? (A3)	Observation
	I have a lot of control over how well I do my job. (AI1)	Interview
	My job allows me to make a lot of decisions on my own. (AI2)	Interview
2. Worker Pace Control (PCS)	How much control does the worker have in setting the pace of his or her own work? (PC1)	Observation
	How much control does the worker himself or herself have over the pace of his/her work? (PC2)	Observation
	The job allows the individual to determine his or her work pace? (PC3)	Observation
	How true is it that the speed at which you work is determined by yourself? (PCI1)	Interview
3. Skill Complexity (SCS)	To what extent does the job require the use of sophisticated or complex skills? (SC1)	Observation
	How intellectually demanding is the job? (SC2)	Observation
	(-)The job is so simple that virtually anybody could handle it, with little or no initial training. (SC3)	Observation
	My job requires a high level of skill. (SCI1)	Interview
	What is the level of school or college you feel is needed by a person doing your job? (SCI2)	Interview
4. Specifiability-Certainty (SPS)	How much uncertainty is there in the job? (SP1)	Observation
	The job requires the individual to be able to handle surprising or unpredictable situations. (SP2)	Observation
	The individual working on the job does tasks which are clearly defined? (SP3)	Observation

Results

The three specific job characteristics used in the study--specifiability, worker pace control, and skill complexity--were analyzed according to the model presented in Figure 1. The correlation[1] represented by A is an indication of the repeatability of the concept being measured by two observers at the same time. The two B correlations provide an indication of the stability of the concept. The λ's represent the relationships of each of the items to the concept being measured.[2] It should be pointed out that the λ's are higher if observer bias exists, but that unless different observers are biased in the same way, the A and B correlations are not.

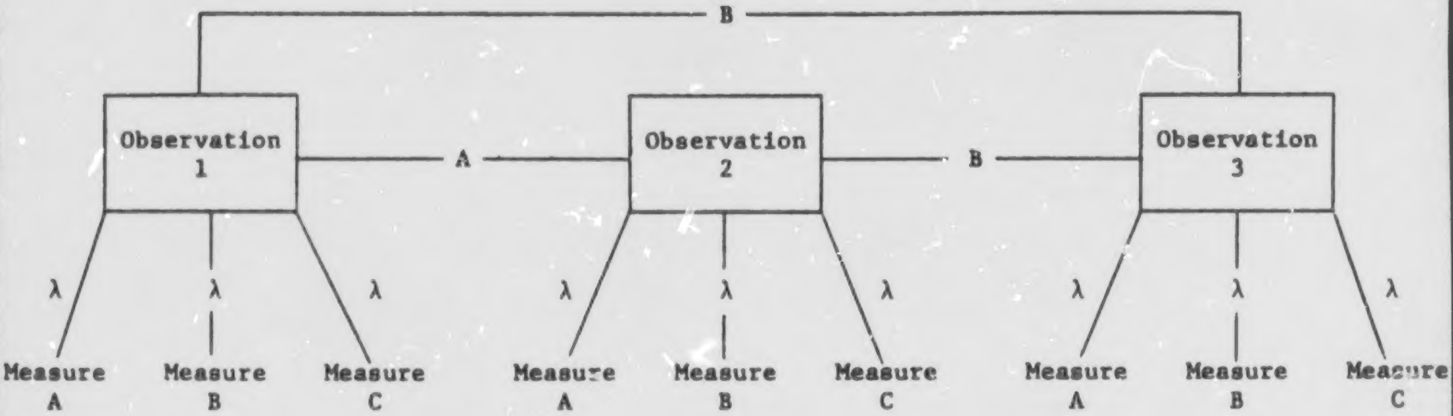
Figures 2, 3, and 4 show the results of these analyses and Tables 2, 3, 4 show the residual matrices indicating the differences between the observed correlations among the measures used in the analyses, and those predicted by the results of the confirmatory factor analysis structure. These analyses indicate:

1. The models in all three analyses provide reasonably good fits of the data. In the model for job specifiability, SP2 is less reliable than the other measures, and the residuals indicate that this measure does not fit the model as well as do the others. As a result, SP2 was dropped from further analyses. Similarly PC2 was dropped from the pace control scale. Excluding these measures, 156 of 168 of the residuals

---

[1] Strictly speaking, these are not correlations, but parameters of the model which are conceptually similar to the correlations among the scales corrected for the unreliability in the scales. They will be referred to as correlations here for ease of presentation.

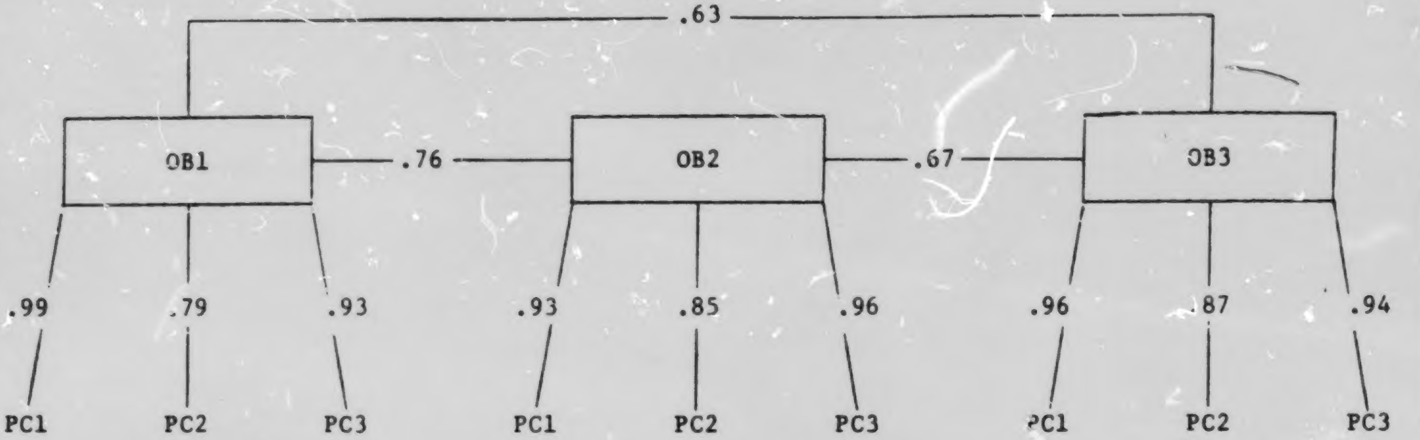
[2] The λ's can be interpreted in a manner similar to factor loadings in a factor analysis which uses an oblique rotation.


B
Observation 1
A
Observation 2
B
Observation 3
λ
λ
λ
λ
λ
λ
λ
λ
λ
Measure A
Measure B
Measure C
Measure A
Measure B
Measure C
Measure A
Measure B
Measure C

Observations 1 and 2 occur at the same time; observation 3 at a different time.

Figure 1

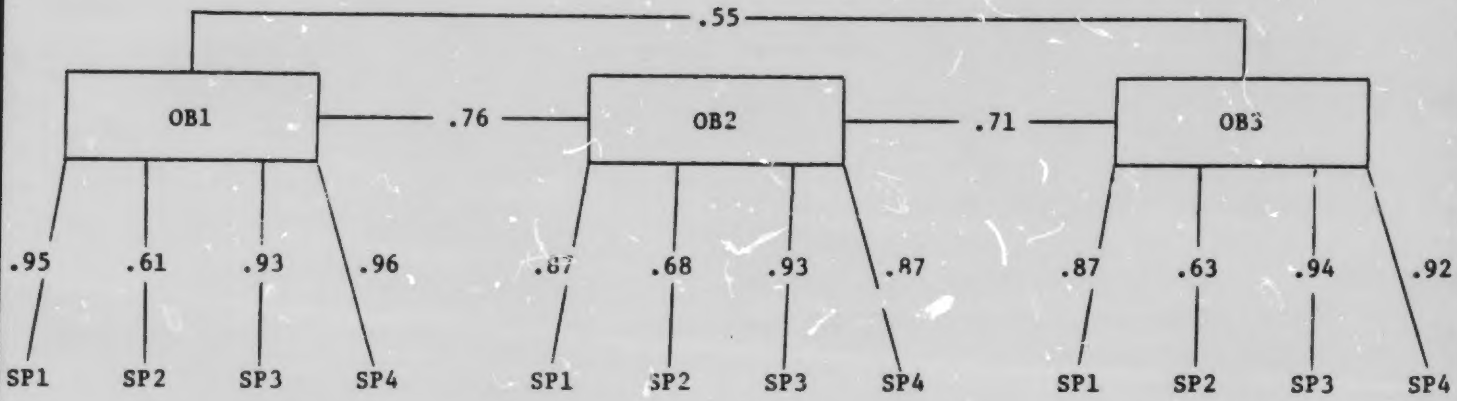
Model for Analyzing the Stability and Repeatability of Job Observation Scales


.63
OB1
.76
OB2
.67
OB3
.99
.79
.93
.93
.85
.96
.96
.87
.94
PC1
PC2
PC3
PC1
PC2
PC3
PC1
PC2
PC3

Observations OB1 and OB2 both occurred at the same time; OB3 at a different time; n = 41; df = 24

Figure 2

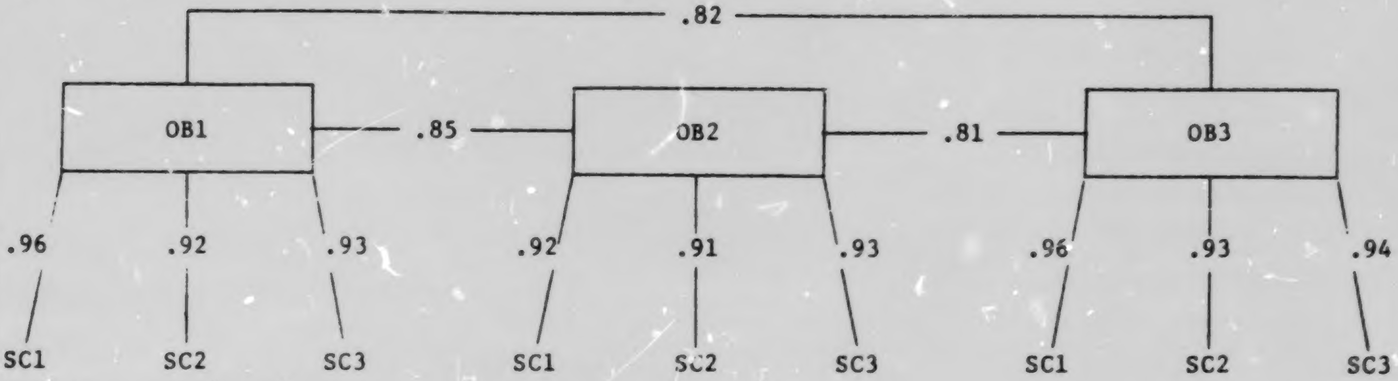
Worker Pace Control


.55
OB1
.76
OB2
.71
OB3
.95
.61
.93
.96
.87
.68
.93
.87
.87
.63
.94
.92
SP1
SP2
SP3
SP4
SP1
SP2
SP3
SP4
SP1
SP2
SP3
SP4

Observations OB1 and OB2 occurred at the same time; OB3 at a different time; n = 41; df = 53

Figure 3

Job Specifiability


.82
OB1
.85
OB2
.81
OB3
.96
.92
.93
.92
.91
.93
.96
.93
.94
SC1
SC2
SC3
SC1
SC2
SC3
SC1
SC2
SC3

Observations OB1 and OB2 occurred at the same time; OB3 at a different time; n = 41; df = 24

Figure 4

Job Skill Complexity

Table 2

Residual Matrix for Pace Control Model

1. OB1-PC1	.00								
2. OB1-PC2	.00	.00							
3. OB1-PC3	.00	.05	.00						
4. OB2-PC1	.02	.08	-.04	.00					
5. OB2-PC2	-.05	-.20	-.07	.02	.00				
6. OB2-PC3	.01	-.02	-.05	-.01	.00	.00			
7. OB3-PC1	.00	.14	-.05	.03	-.10	.07	.00		
8. OB3-PC2	-.05	.06	-.08	.04	-.11	-.01	.00	.00	
9. OB3-PC3	.02	.09	.03	-.05	-.14	-.01	.00	.01	.00
	1	2	3	4	5	6	7	8	9

Table 3

Residual Matrix for Specifiability Analysis

1.	OB1-SP1	.00											
2.	OB1-SP2	-.01	.00										
3.	OB1-SP3	.01	.03	.00									
4.	OB1-SP4	.01	-.03	.00	.00								
5.	OB2-SP1	.02	.08	-.01	-.04	.00							
6.	OB2-SP2	-.02	.19	.06	.06	.01	.00						
7.	OB2-SP3	.03	.08	.02	-.02	-.03	.02	.00					
8.	OB2-SP4	-.02	.13	-.01	-.06	.04	-.10	.02	.00				
9.	OB3-SP1	.11	.14	.09	.08	.22	.15	.15	.17	.00			
10.	OB3-SP2	-.11	.07	-.10	-.12	.07	.14	-.01	-.02	-.04	.00		
11.	OB3-SP3	-.09	.03	-.09	-.04	.00	.01	-.08	-.12	-.01	.01	.00	
12.	OB3-SP4	.00	.07	.06	.08	-.00	.12	-.01	-.10	-.03	-.03	-.03	.00
		1	2	3	4	5	6	7	8	9	10	11	12

Table 4

Residual Matrix for Skill Complexity Analysis

1.	OB1-SC1	.00								
2.	OB1-SC2	-.00	.00							
3.	OB1-SC3	.01	-.02	.00						
4.	OB2-SC1	-.04	.08	.06	.00					
5.	OB2-SC2	-.01	.12	-.03	-.01	.00				
6.	OB2-SC3	-.03	.00	-.01	.02	-.01	.00			
7.	OB3-SC1	.02	.01	.02	.02	.04	-.02	.00		
8.	OB3-SC2	-.03	-.02	-.04	-.07	.06	-.02	.00	.00	
9.	OB3-SC2	.00	.06	-.04	-.05	.08	-.02	-.01	.01	.00
		1	2	3	4	5	6	7	8	9

are less than .10; none of the residuals is significantly different from zero ($p < .05$).

2. In all three analyses, the λ values for the same items were similar for all three observers (shown in Figures 2, 3, 4). This indicates that the structure of the scales is the same for each set of observations, as it should be since the observers were assigned randomly to observations one, two and three.

3. In all three analyses correlation A is greater than the two B correlations, indicating that repeatability is greater than stability for scales. These results also indicate, however, that there is a large degree of agreement among observers at different times (averages of 42%, 40%, and 65% of the variance in the concepts), and that this agreement is not much less than the agreement among observers at the same time (differences average 18%, 16%, and 7% of the variance).

Overall, these analyses support the idea that observers agree in their ratings of these stable job characteristics. The decrease in agreement of observers over time does not seem large enough to be worrisome, and probably indicates that jobs actually are different at different times. Nevertheless, the fact that observers at the same time did not agree with significant frequency (between 23% and 45% of the variance in ratings) suggests the existence of observer bias.

Having shown that the observation scales exhibit reasonable internal consistency, repeatability, and stability, the next step in the analysis was to examine the extent of this halo, and the extent of the ecological relationship among the measures. To this end, the data were analyzed using the models shown in Figures 5 and 6. Both of these models assume that data about jobs were collected by two observers and by an interview with the job holder. The assumption was made in the

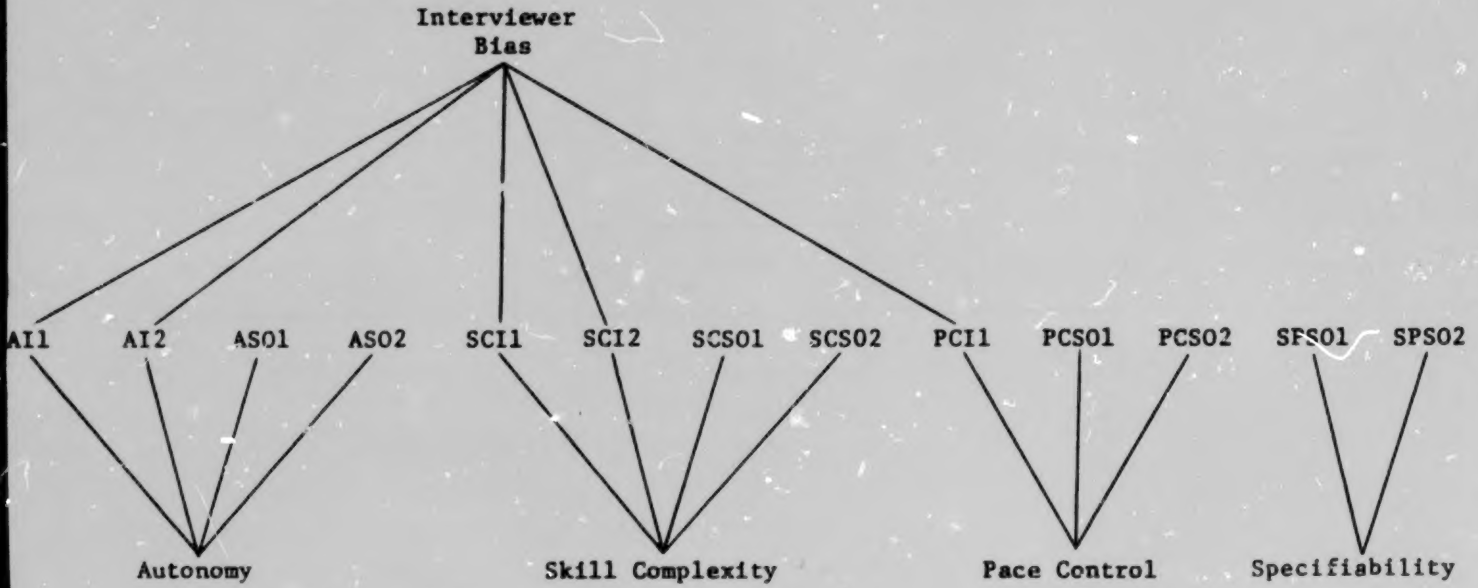

Interviewer Bias
AI1
AI2
ASO1
ASO2
SCI1
SCI2
SCSO1
SCSO2
PCI1
PCSO1
PCSO2
SPSO1
SPSO2
Autonomy
Skill Complexity
Pace Control
Specifiability

Figure 5

Relationships Among Job Characteristics Without Observer Bias

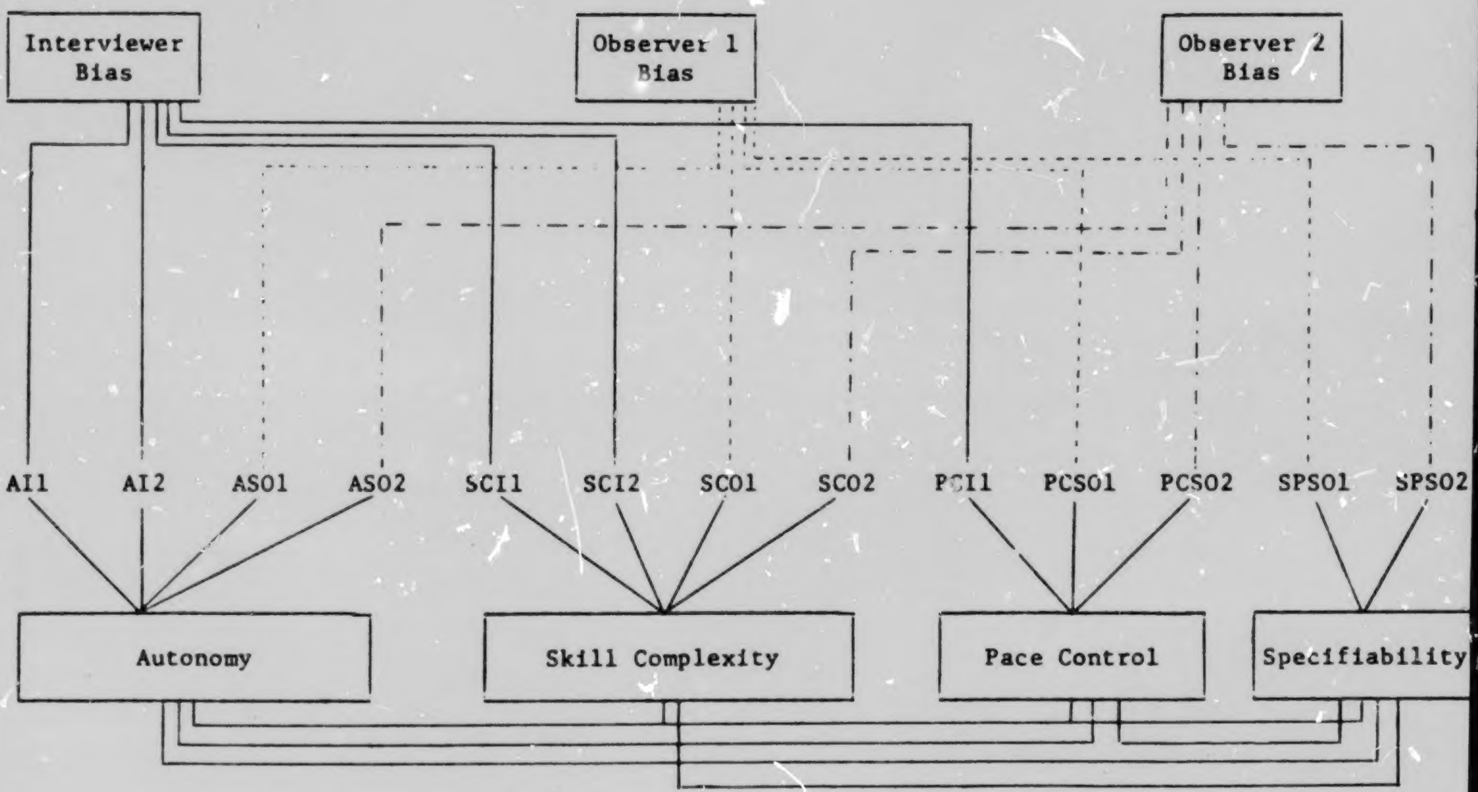

Interviewer Bias
Observer 1 Bias
Observer 2 Bias
AI1
AI2
ASO1
ASO2
SCI1
SCI2
SCO1
SCO2
PCI1
PCSO1
PCSO2
SPSO1
SPSO2
Autonomy
Skill Complexity
Pace Control
Specifiability

Figure 6

Relationships Among Job Characteristics with Observer Bias model shown in Figure 5 that only the interview data were biased, whereas the asusmption in the model shown in Figure 6 is that both the interview and the observation data were biased. In both cases, the effect of observer bias was assumed to be uncorrelated with other bias measures, or with the job characteristics. It was also assumed that since observers were randomly assigned to observation periods, the strength of the biases associated with the observation measures from different observations was the same as were the validities of the different observation measures. The analysis suggested that the model in Figure 6 provided a significantly better fit for the data than the model in Figure 5. It was concluded that observer bias did exist in the data. Only the results of the analysis testing the model in Figure 6 will, therefore, be presented (Table 5). Table 6 shows the residuals resulting from the analysis, and Table 7 the correlations among job characteristics.

The residuals shown in Table 6 indicate that the model provides a reasonably good fit of the data. Table 5 indicates that the observer ratings are biased, but that the effects of the observer bias are not great. The bias effect is greater for the specifiability and autonomy scales than for the skill complexity and the pace control scales, suggesting that the observer halo effect increases as the concreteness of the scale decreases. The same trend appears in the interview data, since the interview bias is smaller for the concrete scales. It should be pointed out that the "bias" indicated for the interview autonomy measure may simply indicate that the job holder's ratings of autonomy reflect information that was not available to the observers (e.g., supervisory style) and does not necessarily reflect invalid autonomy ratings.

Finally, the data presented in Table 7 indicate that even controlling for observer bias, and including interview data, there is still

Table 5

Relationships of Items to Concepts for the Model Shown in Figure 6

Measure	Loading on Scale				Bias Effect		
	Autonomy	Skill Complexity	Pace Control	Specifiability	Interviewer	Observer 1	Observer 2
Aut. Scale - OB1	.802					.511	
Aut. Scale - OB2	.802						.511
Aut. INT 1	.502				.427		
Aut. INT 2	.486				.729		
SC Scale - OB1		.845				.355	
SC Scale - OB2		.845					.355
SC INT 1		.578			.304		
SC INT 2		.779			.018		
Pace - OB1			.728			.344	
Pace - OB2			.728				.344
Pace INT 1			.353		.038		
SP. Scale - OB1				.684		.468	
SP. Scale - OB2				.684			.468

Observers were randomly assigned to observe jobs, so the bias effects and observations of the two observers were assumed to be equal in magnitude. The two observations ocurred at different points in time. N = 532.

Table 6

Residuals Resulting from Analysis of Model Shown in Figure 6

1.	SCSO1	.00												
2.	SCSO2	.00	.00											
3.	SCI1	-.03	.01	.01										
4.	SCI2	.01	-.03	.06	.00									
5.	AI1	-.01	.00	-.01	-.01	.00								
6.	AI2	.01	.04	.01	.00	.00	.00							
7.	PCI1	.03	-.02	-.06	.02	.02	-.01	.00						
8.	PCSO1	-.02	.01	-.01	.01	.00	-.03	-.01	.00					
9.	PCSO2	.03	-.01	.00	-.04	.00	-.04	.02	.01	.00				
10.	ASO1	.00	.02	-.03	.01	.00	.02	.01	.01	.00				
11.	ASO2	.01	-.01	.00	-.04	.01	-.01	-.04	-.02	.03	.01	.00		
12.	SPSO1	.00	-.01	-.04	-.07	-.03	-.01	.00	-.01	.05	-.02	-.01	.00	
13.	SPSO2	-.01	.07	.05	-.02	.06	.08	-.02	.08	.00	.05	.02	.02	.00
		1	2	3	4	5	6	7	8	9	10	11	12	13

Table 7

Correlations Among Job Characteristics Resulting from Analyses of the Model in Figure 6

1. Skill Complexity	1.00			
2. Pace Control	.80	1.00		
3. Specifiability	.91	.70	1.00	
4. Autonomy	.89	.96	.90	1.00
	1	2	3	4

a high degree of covariance among job characteristics. This suggests that the jobs included in this sample were distributed along a single dimension. Jobs high on this dimension were high on skill complexity, worker pace control, and autonomy, and low of specifiability. Thus, the fact that observers do not discriminate among different job characteristics may simply mean that the different characteristics of the jobs were highly related.

Discussion and Conclusion

The analyses indicated that standardized job observation scales are characterized by some degree of observer bias, but that they also exhibit considerable repeatability and stability. Furthermore, the results indicated that previously identified problems of discriminability among observer ratings of different job characteristics may be a function of the fact that job characteristics are ecologically related.

These conclusions must be viewed, however, as tentative. It is possible that the observers, due to the training procedures used, were systematically biased toward rating jobs along a single dimension. If this were true, the effects of this type of bias would affect both observers' ratings, and would produce results similar to those reported here. The fact that the model included interview ratings of the job suggests, however, that this type of bias is accounting for only some covariance among the job dimensions. Thus it seems reasonable to conclude that there is some multicollinearity among job dimensions, but that the results presented here may represent an inflated estimate of its extensiveness.

Overall, the results of these analyses indicate that standardized job observations do represent a useful way of collecting data about jobs. These results also indicate, however, the importance of using multiple methods of data collection whenever possible because each source of data (particularly standardized observations) is biased to some extent, and only when multiple sources of data are used can these biases be estimated and controlled for.

REFERENCES

Alwin, D. An analytic comparison of four approaches to the interpretation of relationships in the multi-trait multi-method matrix. In H. Costner (Ed.), Sociological Methodology 1973-74. San Francisco: Jossey-Bass, 1974.

Joreskog, K., Gruvaeus, G., & Van Thillo, M. ACOVS: A general computer program for analysis of covariance structures. Educational Testing Service (mimeo), Princeton, N.J., 1970.

Werts, C. E., Linn, R. L., & Joreskog, K. G. Quantifying unmeasured variables. In H. Blalock (Ed.), Measurement in the social sciences. Chicago: Aldine, 1974.

Chapter 6

OBTRUSIVE AND LESS OBTRUSIVE MEASURES OF STATUS, SUPERVISORY STYLE, INVOLVEMENT, AND JOB SATISFACTION

by

Terry A. Beehr

ABSTRACT

Obtrusive measures (from interviews), less obtrusive measures (ratings of trained observers), and unobtrusive measures (from company records) were used to measure four important occupational concepts: hierarchical status, supervisory style, job involvement, and job satisfaction. The tests of convergence based on different types of measures generally proved disappointing. A fair degree of convergence was obtained between obtrusive and other measures of hierarchical status. No convergence emerged between the two types of measures of supervisory style. As regards job involvement, the less obtrusive (observational) measures did not intercorrelate substantially, let alone correlate with the obtrusive (interview) measure. In the case of job satisfaction, intercorrelations were appropriately positive among the obtrusive (interview) measures but only marginally positive among the less obtrusive (observation) measures. Correlations between the two types of measures yielded only minimal convergence, and other relevant correlations suggested that observers' estimates of job satisfaction were based on the degree of challenge the job offered rather than the job occupant's response to the job. By way of interpretation, various reasons were cited for both the successful and unsuccessful attempts to achieve convergence between obtrusive (interview) and less obtrusive modes of measurement.

Chapter 6

OBTRUSIVE AND LESS OBTRUSIVE MEASURES OF STATUS, SUPERVISORY STYLE, INVOLVEMENT, AND JOB SATISFACTION

Obtrusive measures such as questions asked in interviews have at least two inherent problems: (1) subjects may be motivated to provide inaccurate data; and (2) interviews tap the perceptions of only one person--the respondent. Measures other than questions in interviews can, if unobtrusive, minimize and sometimes circumvent these problems. Yet obtrusiveness is a continuous rather than a dichotomous variable, and the attempt in this study was to employ two methods less obtrusive than questions in interviews.

The strategy was to assess the degree of convergence between the less obtrusive measures and the obtrusive (interview) measures of the same phenomena (and to assess the degree of discrimination among measures of different phenomena) obtained from each method.

Three modes of data collection were employed in the present study. In addition to (obtrusive) interview measures, the study included the observations of trained observers (less obtrusive than interviews because the workers were attending to their work as well as to the presence of observers) and data from company records (unobtrusive). As many modes of data collection as possible were used to assess four variables:

hierarchical status, supervisory style, job involvement, and job satisfaction.

Method

Sample

This report concerns only Phase I of this study and includes data on employees from five organizations: a hospital, a printing company, a research and development laboratory, and two plants that manufactured automobile accessories--all in the midwest. Though not strictly a sample of a specified population, the respondents were drawn from a variety of occupations. Compared to a national population the analysis sample contains a disproportionately large number of female workers, young workers, blacks, single people, and operatives. A more extensive description of sampling procedures and sample characteristics is provided in Chapter 1.

Measures

A detailed account of the three modes of data collection appears in Chapters 1, 3 and 4 in this volume. Briefly, professional interviewers conducted personal interviews in the homes of employees, each interview lasting about two hours. Each respondent was also observed on the job for an hour on two separate occasions by two different observers who had previously been trained to rate a variety of characteristics of worker and job. In addition, organizational records were made available by management to the research staff to code such basic file information as respondent's annual income.

Results and Discussion

Status

Five variables expected to vary as a function of hierarchical status in organizations were measured: mental skill level required to do the work, status symbols (e.g., carpets on the floor or drapes on the windows), the amount of freedom allowed in doing one's work, income, and being a supervisor of other employees. An inspection of Table 1 shows which of these variables were measured through interviews, observations, and company records.

Table 1 shows the correlations among the eight status variables. The variables all had one thing in common, viz., they were expected to vary with hierarchical status. All of the correlations were positive and significant ($p < .05$). The correlations among the four measures less obtrusive than interviews (upper left triangle) were quite strong, indicating that they did have a common element. The correlation between mental skill level and freedom in that triangle was so strong (.81) that it could be inferred that the observers did not distinguish between these two variables. That the concepts were indeed separate was indicated by the relatively lower correlation (.30) between them in the interview data (lower right triangle).

The correlations among the four interview variables were lower than the correlations among the three observation variables, indicating that, while the variables had a common element (status), the respondents were able to distinguish among them better than were the observers.

Table 1

Correlations Among Aspects of Status[a] (N=460)

	1	2	3	4	5	6	7	8
LESS OBTRUSIVE								
Observations								
1. Mental skill level								
2. Status symbols	.54							
3. Amount of freedom	.81	.45						
UNOBTRUSIVE								
Records								
4. Annual income	.48	.38	.39					
OBTRUSIVE								
Interviews								
5. Mental skill level	.67	.43	.58	.41				
6. Supervisory status	.52	.35	.51	.47	.47			
7. Enough freedom	.22	.15	.24	.13	.30	.13		
8. Annual income	.44	.44	.34	.71	.36	.36	.15	

[a] All correlations were significant beyond the .01 level.

The correlations between the obtrusively and the less obtrusively gathered data are also in Table 1. Since each variable measured only one aspect of status, the correlations were expected to be significant but less than perfect. Overall, this expectation was confirmed; most of the correlations were of medium strength.

Identical Concepts Measured by Different Methods. Three of the specific aspects of status, viz., mental skill level, freedom, and annual income were measured by interview and by a less obtrusive method. The variable measuring each of these aspects from one source should have corresponded more closely to the variable from the other source measuring the same aspect than it corresponded to any other variable in the matrix.

In general, the interview measures met this test. The only exception was the interview variable measuring freedom, which correlated more strongly (r = .30) with the interview measure of mental skill level than with the observation measure of freedom (r = .24).

The less obtrusive measures also met this test with one exception: It must be conceded that the observers were not able to distinguish between a job's freedom and the mental skill level required, since observed freedom was more highly correlated with mental skill level as measured by either method than with freedom as measured by the interview. If observed freedom were eliminated from the matrix, observed mental skill level was more strongly correlated (r = .67) with the interview measure of mental skill level than with any other variable in the matrix. Also, recorded annual income was more strongly correlated with the interview measure of annual income than with any other variable.

Analysis by Organizational Site. It was possible that some of the measures of status were more valid within a single organization than across several organizations. If this were true, the organization would act as a suppressor variable, i.e., the correlations for the whole sample would be smaller than the correlations for any single subsample. Inspection of correlations among the status variables for subjects in each of the five sites (not shown here) did not support this premise, however. No overall trend toward larger correlations within sites than for the whole sample was detected.

It was concluded that less obtrusive measures of variables indicating status were obtainable from both company records and standardized observations.

Supervisory Style

Four supervisory style variables were measured through the interview: psychological support offered by the supervisor to his/her subordinates, the supervisor having high work standards for himself/herself, the amount of feedback given to subordinates, and the extent to which the supervisor insisted on hard work from subordinates. One of these, supervisor feedback, was also measured by observation. In addition, observers rated the physical proximity of the supervisor to his/her subordinates, the degree to which the details of a subordinate's job were closely supervised, and the degree to which the supervisor exerted unreasonable pressure for production. It was expected that there would be a moderate relationship between these last three observation variables and the interview measures of high work standards for self and insisting on hard work from subordinates.

Overall, the attempt to measure supervisory style with less obtrusive measures yielded dismal results. None of the correlations between observation and interview supervisory style measures was significant beyond the .05 confidence level (Table 2). Apparently, measures obtained by the two methods were not measuring the same things at all. Even the supervisor feedback measures, two virtually identical items measured by different methods, were not related to each other.

Relationships Among the Observed Variables. All of the correlations among the observation measures were positive, significant, and of low or moderate strength. They are about as strong as expected, given that the variables in the triangle were probably related but not identical.

Relationships Among the Interview Variables. The relationships among the interview variables were all positive and about as strong as the relationships found among the observed variables. It was puzzling that a few of them were so strong, e.g., the correlation between supervisor support and supervisors having high work standards. If it were not for the low correlations between the supervisor insisting on hard work and the other variables, it might be concluded that a halo effect or a response bias accounted for much of the strength of the relationships. Those low correlations repudiated this interpretation, however.

Since the observers spent very little time with the supervisor relative to the time that the interview respondent spent with the supervisor, the interview data were probably more accurate than the observer ratings.

Table 2

Correlations Among Different Types of Supervisory Style
(N=439)

	1	2	3	4	5	6	7	8
LESS OBTRUSIVE								
Observation								
1. Proximity								
2. Close supervision	.55**							
3. Unreasonable pressure	.18**	.14**						
4. Supervisor feedback	.53**	.54**	.38**					
OBTRUSIVE								
Interview								
5. Support	.00	-.01	-.06	-.04				
6. High work standards for self	-.09	-.09	-.02	-.04	.62**			
7. Supervisor feedback	-.03	-.07	.01	-.04	.60**	.49**		
8. Insists on hard work	-.03	-.03	.06	-.01	.07	.20**	.07	

*p < .05

**p < .01

Also, supervisory behavior was likely to have been fairly reactive to the presence of an observer. It is possible, therefore, that the observers were not as unobtrusive as they might have been, and that their presence affected the supervisor's behavior.

Job Involvement

Because job involvement is an intrapsychic phenomenon, it was expected that the best measure would be self-report, i.e., the interview. Therefore, for this analysis, the strategy was to validate the observer's ratings against the interview measure of involvement.

One interview item asked respondents how involved they were in their work compared to their other interests. Observers rated the employees' involvement directly, and they also rated the frequency of some employee behaviors that could indicate affect toward the work. These behaviors were combined into two categories: laughing/smiling and aggressive behaviors.

Only one correlation in Table 3 is significant, and it is very low. Neither the employees' laughing/smiling behaviors nor their aggressive behaviors were good indicators of involvement. It is noteworthy that the observed measures were not related even to each other. Apparently, when the observers rated involvement, they did not use observed laughing/smiling or aggressive behaviors as clues. Nevertheless, _whatever_ the clues to which the observers were attending, they were not good indicators of job involvement, since the correlation between observed involvement and interview involvement was not significant.

There remains the question of whether the observers _could_ have predicted

Table 3

Correlations Among Indicators of Involvement
(N=574)

	1	2	3	4
LESS OBTRUSIVE				
Observation				
1. Involvement				
2. Laughing/smiling	.10*			
3. Aggression	.05	.06		
OBTRUSIVE				
Interview				
4. Involvement	.01	.03	.04	

*p < .01

employees' involvement. Hackman and Lawler (1971), following a trend of research (Turner & Lawrence, 1965; Blood & Hulin, 1967) on four task characteristics (variety, autonomy, task identity, and feedback), found that these characteristics were significantly related to employees' intrinsic motivation and job involvement. Employees performing tasks that are experienced as having these characteristics tend to be strongly involved in their work.

The observers in the present study rated the employees' jobs on the core characteristics (Chapter 4). It was possible that the observers could have estimated the employee's involvement from the characteristics on his job. This approach would assume that job involvement was, at least in part, a function of the task. Table 4 indicates the extent to which observers could have inferred the involvement of employees from observing their job characteristics.

Since all four of the observed core characteristics were correlated significantly with the interview involvement measure, the observers could have estimated job involvement from their observations of the task. It is clear that they did not do this, however, since their ratings of three of the four task characteristics were unrelated to their ratings of involvement, and their rating of the fourth task characteristic, i.e., task identity, had only a small correlation with observed involvement.

It is concluded that the unobtrusive measures of involvement were probably invalid. Observers could have inferred employees' involvement from task characteristics, but this would not have been a direct measure of involvement.

Table 4

Correlations Between Observed Task Characteristics and the Job Involvement Indicators (N=574)

Observed Task Characteristics	Involvement Indicators			
		Observation		
	Interview Involvement	Involvement	Laughing/ Smiling	Aggression
Variety	.43**	.04	.14**	.13**
Autonomy	.45**	.05	.15**	.16**
Task identity	.26**	.11**	.09*	.08*
Task feedback	.12**	-.02	.05	.07

*p < .05

**p < .01

Job Satisfaction

Four possible measures of overall job satisfaction were included in the observation schedule: (1) laughing/smiling behavior and (2) aggressive behavior of the respondent, (3) a direct rating of the respondent's satisfaction, and (4) a rating of the degree to which the observers thought they would feel satisfied if they had the job. The first two were behaviors thought to be indicative of affect and were therefore tested as possible measures of both of the affective phenomena in this study, i.e., involvement and satisfaction.

Because satisfaction (like involvement) is an intrapsychic phenomenon, it was assumed that the interview measures would be more accurate than those provided by observers.

The interview asked respondents to rate their satisfaction with four facets of the job: challenge, comfort, resources, and financial rewards. There was also an interview measure of overall, facet-free job satisfaction.

In addition to these satisfaction measures, there were measures of the job facets themselves. Respondents were asked to rate the degree to which their job offered challenge, comfort, resources, and financial rewards (Barnowe, Mangione & Quinn, 1972). Observers rated the jobs on challenge, comfort, and resources, and a measure of financial rewards was obtained from the company records.

Table 5 shows the correlations among measures of satisfaction obtained from observations and interviews. Most of the observation measures were not highly correlated with each other. The major exception was the correlation (.74) between the observer's estimate of the respondent's satisfaction and the observer's estimate of the degree of satisfaction

Table 5

Correlations Among Indicators of Job Satisfaction (N=574)

	1	2	3	4	5	6	7	8	9
LESS OBTRUSIVE									
Observation Satisfaction									
1. Laughing/ smiling									
2. Aggression	.07								
3. Respondent's satisfaction	.19**	.03							
4. Observer's satisfaction	.14**	.09*	.74**						
OBTRUSIVE									
Interview Satisfaction									
5. Challenge	.01	-.01	.30**	.21**					
6. Comfort	.02	-.05	.13**	.07	.64**				
7. Resources	-.04	-.09*	.11**	.01	.63**	.66**			
8. Financial rewards	-.03	-.07	.17**	.13**	.49**	.48**	.48**		
9. Facet-free	-.01	.02	.20**	.10**	.56**	.41**	.44**	.39**	

*p < .05
**p < .01

that he/she would feel if he/she were doing the job himself/herself.

This strong correlation had at least three possible implications. First, it might indicate that observers were projecting, i.e., the observers thought the respondents were satisfied to the same extent as they would have been, because each observer tended to see the respondents as similar to himself/herself. Second, a job might affect most people in a similar way, i.e., the observer and the respondent would react in a similar fashion to the same stimuli. Third, observers may have seen that the respondent was satisfied and concluded, therefore, that they probably would have been satisfied with that job also.

Observed laughing/smiling was correlated significantly with observed respondent's satisfaction (r = .19) and with observed observer's satisfaction (r = .14), although the correlations were low.

The correlations in the lower right corner are all fairly strong and indicate that the interview measures of satisfaction measured some common entity. The correlations among the five types of satisfactions were higher than might be expected, perhaps indicating that the respondents were experiencing a halo effect. i.e., their satisfaction with one facet influenced their satisfaction with other facets and their general satisfaction with the job.

In general, there was very little convergence between the two methods. Only two observation measures, observed respondent's satisfaction and observed observer's satisfaction, showed any consistent convergence with the interview satisfaction measures. The correlations between these two observation satisfaction measures and the interview satisfaction measures were somewhat low, but two interesting patterns were present.

Patterns Within the Matrix. First, observed respondent satisfaction was correlated more highly with each interview satisfaction measure than observed observer satisfaction was. This was encouraging, since it indicated that observers were consistently able to distinguish, to some extent, between their own reactions to the job and the respondent's reactions to it. Although the differences were not large, they were all in the expected direction.

Second, it was expected that the observation measures of satisfaction would measure overall or facet-free satisfaction. Instead, the interview variable that was correlated most strongly with both observed respondent satisfaction and observed observer satisfaction was interview challenge satisfaction (.30 and .21, respectively). The next strongest correlates of observed respondent satisfaction and observed observer satisfaction were facet-free satisfaction and financial rewards satisfaction, respectively.

Since the strongest correlations with the measures of observed satisfaction were with interview challenge satisfaction, it was possible that observers were focusing on aspects of the job related to perceived challenge. This was particularly likely because interview challenge satisfaction was related strongly to interview satisfaction (r=.56).

In addition, Table 6 shows that the two observation satisfaction measures were related more strongly (1) to interview challenge than to any other interview job facet measure and (2) to observed job challenge than to any other observation job facet. This was good evidence that observers were focusing on one job facet, viz., challenge, more than others in rating satisfaction.

The greatest insight into the observers' criteria for rating job satisfaction was obtained by comparing the correlations in Table 5 with the correlations in Table 6. Specifically, both of the observed satisfaction measures in Table 6 correlated more strongly with the less obtrusive measures of two job facets (challenge and financial rewards) than with any interview satisfaction measure in Table 5. It seems that observers were rating satisfaction on the basis of the quality of those two job facets rather than on the basis of the employees' behaviors.

The strength of the correlations between the observation satisfaction measures and observation challenge indicated clearly that the observers ranked jobs almost identically on satisfaction and challenge. Their perceptions of challenge seem to have had a very strong impact on their ratings of both of their measures of satisfaction.

Obviously, the observers did not know what the financial rewards for particular jobs were, but there are two plausible explanations for these high correlations. First, observers may have had a stereotypic but relatively accurate conception of a hierarchy of jobs in terms of financial rewards. Second, financial rewards may have been correlated strongly with job challenge, accounting for the correlations between measures of

Table 6

Correlations of Observed Respondent's Satisfaction and Observed Observer's Satisfaction with Job Facet
(N=545)

Job Facets	Correlations with Observation Satisfaction	
	Respondent's Satisfaction	Observer's Satisfaction
Interview		
Challenge	.44**	.38**
Comfort	.08*	.02
Resources	.05	-.02
Financial Rewards	.30**	.36**
Observation		
Challenge	.75**	.81**
Comfort	.21**	.26**
Resources	-.10*	-.12**
Records		
Financial Rewards	.38*	.39**

*p < .05

**p < .01

financial rewards from the two sources and the observation measures of satisfaction. Observers' ratings of challenge were strongly related to recorded financial rewards (r = .56) and moderately related to interview financial rewards (r = .33) [not shown in tables]. This pattern of correlations matched the pattern in Table 6, i.e., the recorded measure of financial rewards was related to the observed satisfaction measures somewhat more strongly than was the interview measure of financial rewards.

Thus, the fact that the observation satisfaction measures were related to interview satisfaction (Table 5) was probably due primarily to the relationship of each of these measures to job facets (Table 6). The relationship between observation satisfaction and interview satisfaction was not direct. Instead of measuring satisfaction, the observers appeared to be inferring satisfaction from observed job characteristics, especially challenge.

Summary and Conclusions

Only very limited success may be claimed for the less obtrusive measures tested against obtrusive measures in this study. Not surprisingly, the less obtrusive measures worked best for concepts with relatively visible and explicit indicators, such as hierarchical status, although even here the degree of agreement between obtrusive and less obtrusive measures was insufficient to warrant treating them as substitutable. On more subtle dimensions such as supervisory style there was substantial agreement among obtrusive measures and among unobtrusive measures but no relationship between the two classes of measures. Perhaps longer periods of observation are needed to ensure reliable ratings of interpersonal phenomena. Even more dismal results greeted job involvement, the first of the intrapsychic concepts. There was only minimal agreement among the less obtrusive measures and no association between them and the interview measure of involvement. While it was obviously hard for observers to rate intrapsychic phenomena reliably, the behavior singled out as potential indicators of involvement (laughing/smiling and aggression) tapped affect in general rather than job involvement in particular and one of them (laughing/smiling) could even be thought to indicate low involvement in the sense of "not being serious." The data on job satisfaction, the second intrapsychic concept, fared only slightly better. The interview measures did intercorrelate fairly strongly, but with one exception (respondent's satisfaction and observer's satisfaction) the observation measures did not. The quite limited correlations between the two types of measures of satisfaction suggested that, when rating satisfaction, observers were responding primarily to the actual challenge offered by the job rather than to the respondents' reactions to a series of job facets.

Clearly, this study did not produce measures of other than obvious or visible occupational variables that were both reliable and less obtrusive. Whether this can be done with better selections of respondent behaviors on which observers can key, longer periods of observations, or improved interview measures as comparison points remains to be seen. Hopefully it can, thereby leading to the strategy recommended by Webb, Campbell, Schwartz, &

Sechrest (1966) of combining measures from several sources. Such a strategy would have the advantage that one measurement source may be strong where others are weak and that the combination of several methods may surpass the value of any single measure.

References

Barnowe, J. T., Mangione, T. W. and Quinn, R. P. The relative importance of job facets as indicated by an empirically derived model of job satisfaction. Ann Arbor, Mi.: Survey Research Center, 1972 (mimeo).

Blood, M. R. and Hulin, C. L. Alienation, environmental characteristics and worker responses. *Journal of Applied Psychology*, 1967, *51*, 284-290.

Hackman, J. R. and Lawler, E. E. III. Employee reactions to job characteristics. *Journal of Applied Psychology Monograph*, 1971, *55*, 259-286.

Turner, A. N. and Lawrence, P. R. *Industrial jobs and the worker*. Boston: Harvard University Graduate School of Business Administration, 1965.

Webb, E. J., Campbell, D. T., Schwartz, R. D., and Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally & Company, 1966.

Chapter 7

STANDARDIZED OBSERVATIONS OF JOB CHARACTERISTICS: A REFINEMENT AND REPLICATION

by

G. Douglas Jenkins, Jr. and David A. Nadler

ABSTRACT

A replication and refinement of the standardized job observation technique was attempted. Training procedures and instrument changes were employed to address the deficiencies of the previous study. In all, 147 employees were observed by 10 trained observers. Off-the-job interview measures were also obtained. Observation measures of job characteristics were assessed as to the repeatability of the operations, homogeneity of scales and convergence and discrimination of the constructs measured by observations with constructs measured by employee interviews. Results substantially replicated the findings of the first study with respect to repeatability and homogeneity of measures. Results indicated improvement from the first study in convergence and discrimination of the observation scales. Cognitive social psychological theories are proposed to account for the results.

Chapter 7

STANDARDIZED OBSERVATIONS OF JOB CHARACTERISTICS: A REFINEMENT AND REPLICATION

Introduction

Concerns about the quality of working life as well as employee motivation and productivity have led to new efforts to identify the attributes of jobs and work tasks that have psychological implications for the job holder. Recent work (Hackman & Oldham, 1975, 1976) substantiates the possibility of identifying specific characteristics of work that are measurable, are generalizable across jobs, and are associated with psychological states of the job holder.

In an earlier study (Jenkins, Nadler, Lawler, & Cammann, 1975; also Chapter 4) an attempt was made to develop a standardized method for observing the characteristics of jobs. These authors argued that total reliance on self-report measures was problematic, and that significant differences in job characteristics, if they exist, should be observable. In the earlier study a standardized job observation guide was constructed based on the Hackman and Lawler (1971) job characteristics framework. A variety of methods were employed, including the use of Likert-type and behaviorally anchored rating scales. Observers were trained in a two-day session using videotaped jobs, ratings of these jobs, and feedback on ratings as the primary training tool. The observers used the job observation guide to make 914 one hour observations in three different organizations on 448 jobs. At the same time, the job holders also underwent a two hour structured interview on work roles, working conditions, job characteristics and related issues (details in Chapter 1). The stability, homogeneity, and convergence of the observational measures were then tested with both multiple observation data and data from the interviews.

The present paper reports further work on the standardized observation technology. Based on the findings of the Jenkins et al. (1975) study, changes were made in the training process, the observation instrument, and the associated interview. According to the earlier analyses, these changes should have increased the effectiveness of this method and, thus, its utility for a variety of scientific and applied purposes. The current study was thus designed to test whether these modifications would indeed result in a more reliable and valid measurement technology.

Drawbacks of the Initial Technology

The results of the initial study were moderately encouraging. Three different types of analysis were used to assess the effectiveness of the job observation instrument and process. The first analysis tested the stability of the instrument and process. It asked whether two trained observers using the instrument and observing the same work (either at the same or different time periods) would arrive at the same job ratings independently. Analyses using Cohen's weighted kappa (a chance corrected measure of nominal scale agreement, Cohen, 1968) showed generally high agreement among observers. A second analysis tested the internal consistency of the multiple item scales in the instruments. Analysis focused on whether items designed to tap the same job dimension did indeed correlate. Here the findings were encouraging, with moderate to high scale reliabilities for most of the major job dimensions. The final analysis assessed the convergent validity of the job observation technology by determining the extent to which job observations converged with self-reports on the same job. A multitrait-multimethod matrix was constructed. Only moderate levels of convergence were attained and the degree of discrimination among different job characteristics by observers was disappointingly low. At the conclusion of the initial study two reasons were suggested for the somewhat disappointing results:

Instrument problems. The reporting instrument used was exceptionally long and required the observation of a large range of activities. Some of the data sought were not central to the testing of the observational method and indeed did not prove useful. These portions of the instrument added a clerical and attention burden to the observer, adding to observer fatigue. Some of these portions (e.g., the counting of interactions and the counting of complete work cycles) consumed much time the observer could otherwise have given to more central tasks of observing and rating the job. Finally, a lack of consistency of format and wording between the interview and the observation items may also have contributed to the low convergence scores.

Training. The observer training program focussed on agreement among raters, but did not deal with the issue of agreement with any standard or external norm. The observers were trained to be consistent with each other, but nothing in the program precluded their being consistently wrong in their ratings.

Instrument Changes

For the subsequent study reported here, a number of critical changes were made in the observers' guidelines and reporting instrument, based on the findings and suggestions from the initial observation work. The specific changes were as follows:

Instrument length. All non-necessary items were removed from the instrument as were most of the operations which did not appear to "work" in the first study. This substantially shortened the instrument and removed a whole set of tasks from the observer.

Focussing questions. Certain observer questions and procedures were redesigned to help the observer focus his/her attention initially on the critical elements of the job before the longer period of detailed observation and before the actual rating of the job by the observer. These questions involved working conditions, sources of information and materials, types of equipment or tools used, and the length of time between repeated cycles of work activity.

Interpersonal activity codes. As part of these preliminary observations the observers were asked to code the people-related activities of the job holder, using modified descriptions from the Dictionary of Occupational Titles (U.S. Department of Labor, 1965).

Formatting. The instrument was reformatted to be more readable and easier to use in the field.

Consistency. Changes were made so that more items would be consistent between the observation guide and the interview.

Changes in Observer Training Program

Conceptual materials. More time was taken in training to explain to the observers what they were observing and why. Observers received a lecture on basic concepts of job design and were given a sheet with definitions of critical key terms and concepts. They were thus oriented in a general sense to be able to discriminate between different types of job characteristics.

Development and use of expert ratings of tapes. In order to move beyond simple agreement between observers, an attempt was made to develop a standard against which the observers could test their observations. For each video-taped job, a panel of expert raters (researchers at the Institute for Social Research) used the job instrument to rate the job. The ratings were discussed, modifications made if necessary, and explanations recorded for the ratings. These were then summarized as the expert ratings. During training, instead of just comparing ratings to each other, the observer-trainees' ratings were compared to the expert ratings and feedback was given as to the rationale for the expert ratings. Thus, some standard of validity of ratings was used, both for training and for the selection of observers.

Wider range of tapes. New video tapes were developed encompassing a wider range of job types. Efforts were made to get a broad range of jobs as well as jobs that were high on some job dimensions while low on others, providing an opportunity to test and train the observers to discriminate among the different dimensions.

These changes were intended to exploit what Kerlinger (1964, p. 505) describes as the strength of observation techniques: ". . . the observer can relate the observed behavior to the constructs or variables in the study: he brings behavior and construct together."

A copy of the revised job observation instrument appears in Volume II, Appendix G, along with a copy of the schedule and conceptual materials from the observer training program.

Methods

Subjects

The overall design of the two-phase study of which this investigation is a part is described in Chapters 1 and 2. The present report is based upon data from 147 of the 272 employees who participated in the second phase of the study. Observations were made at only two of the three organizations where respondents were interviewed: the automobile parts manufacturer, and the three departments of the hospital. One of the organizations was unionized. The respondents in the sample represented a wide variety of jobs. Using Census Occupational Classifications (U.S. Bureau of the Census, 1971), most of the jobs fell into six major categories: operative (19.2%), clerical (12.3%), professional/technical (26.0%), service (20.5%), craftsmen/foremen (6.2%), and managers/officials (15.1%). Fifty-five percent of the subjects earned over $10,000 and 57.8% had completed high school. Sixty-one percent of the subjects were white and 47% were male. Each respondent was told at the conclusion of the interview that the observations were planned, and emphasized that the observer was studying working conditions, not how hard or how well the employee worked. The subjects were then given the option of not being observed, but few declined.

Selection and Training of Observers

Announcements were posted stating that observers were wanted for a research project. Applicants were told that they would have to attend a two-day training session before they were hired, but that they would be paid for attending the session. In the training session, the applicants worked on simulated observational tasks and obtained feedback on their performance. This method enabled the training activities to be used for selection purposes, since trainee performance on the simulated observation task could be used as an objective measure and, thus, predictor of observer performance on the job.

Four jobs similar to those performed by employees in the sample were video-taped prior to the training. During the two-day training period each trainee rated the video-taped jobs using the observation instrument. Each rating period was followed by a session in which the trainees compared and discussed their ratings. The training staff encouraged discussion where there were important differences among the ratings of the trainees. Individuals talked about why they rated the tapes as they did, what cues they attended to, and how they reached a decision for each rating. In contrast to the first phase of the study in which the stated goal of the sessions was to move toward agreement among observers as each iteration of the cycle occurred, pooled judgments of experts were used in the second phase as the standard observers were to approach. These ratings were by four members of the research staff knowledgeable in the area of job and task design.

The ratings on the final observation of tapes were collected and used to select observers. Analyses of scores using three different measures of deviation and agreement were performed, and permanent hiring decisions were then made. Of the 18 trainees, 10 were hired.

The observers were non-professionals. Almost all were college students, most of them juniors, seniors, or first year graduate students. The group included five men and five women. They were paid $2.90 per hour and most worked on a part-time basis.

Observation Instrument

The observation instrument included 38 job description questionnaire-type items measuring diverse job factors. The items are presented in the stubs of Table 1 grouped according to their *a priori* scale assignments. The majority of these items were responded to on 4- and 6-point Likert-type scales; the remainder on 7-point anchored scales (see Hackman & Lawler, 1971; Jenkins et al., 1975). In addition, an attempt was made to have observers code the "people-related" activities of the job holder, using modified descriptions from the *Dictionary of Occupational Titles*, (U.S. Department of Labor, 1965), and to code the extent of the interactions the job holder had with specified "others".

Observation Procedures

Each job/respondent in the sample was observed at least once by a single observer for approximately 30 minutes. A randomly drawn sample of approximately one-third of all jobs were also observed on a separate occasion with two observers present and observing at the same time. The joint observations were made by observers other than the one who observed the job alone. The ordering of the "joint" and "alone" observations was arbitrary. Such a design permits assessment of the repeatability of each measure when observers were viewing the same job holder at either the same or at different times. In all, there were 166 hours of observation. Observers were instructed to structure their observation period so that

5-15 minutes were spent becoming oriented to the job, 15-45 minutes in general observation of the job, 15 minutes observing specific job actions, and 15 minutes rating the job *in situ*. A few minutes were usually spent "cleaning-up" the observation instrument away from the job setting.

Analytic Procedures

Several measures of inter-observer agreement or repeatability for nominal and ordinal scales exist in the psychometric literature. Following Jenkins et al. (1975), Cohen's weighted kappa ($\kappa_{\underline{w}}$) (Cohen, 1968; Fleiss, Cohen & Everitt, 1969) was selected as the measure of repeatability. It is a chance corrected estimate of the proportion of agreement between raters. In addition, the statistic allows for partial as well as full agreement by assigning differential penalties to deviations from perfect agreement. Because of the ordinal nature of the scoring system when there was perfect agreement among observers (i.e., Observer A rated 3 and Observer B rated 3), full agreement credit (or a weight of 1.0) would be given. When the ratings differed by only one category (i.e., Observer A scored 3 and Observer B scored 2 or 4), then one-half agreement credit would be given (or a weight of 0.5). Disagreements of more than one category between observers were given no agreement credit (or a weight of 0.0).

In general, $\kappa_{\underline{w}}$ is computed by subtracting the weighted proportion of agreement expected by chance ($\underline{p}_{\underline{c}}$, the sum of the weighted cross product of the marginals) from the weighted observed proportion of agreement and dividing by the maximum chance-corrected agreement possible ($1 - \underline{p}_{\underline{c}}$).*

---

*For the complete formulae and computation procedures of $\kappa_{\underline{w}}$ and its variance estimates see Fleiss et al. (1969).

Theoretically, $\kappa_{\underline{w}}$ may range in value from -1 to +1; however, maximum values require identical marginals. The extent to which the marginals differ lowers the practical maximum value of $\kappa_{\underline{w}}$ (Cohen, 1960).

A $\kappa_{\underline{w}}$ value of 0 would indicate no increase in agreement above the "chance" level estimated from the cross product of marginal proportions. Negative values of $\kappa_{\underline{w}}$ represent levels of agreement below what would be expected by chance alone. Positive values represent levels of agreement between observers over what would be expected by chance.*

Homogeneity

The most common measure of the homogeneity or internal consistency of a set of items which comprise a scale designed to be homogeneous is Cronbach's coefficient alpha (the generalized equivalent of Kuder-Richardson Formula 20). But because a major determinant of the value of this coefficient is the number of items in the scale, the coefficient can be very high, even in cases of low covariation among items. While the scales involved in the observation instrument are relatively short (none exceed five items) it is still preferable to assess the homogeneity of the

---

*Care should be taken not to confuse $\kappa_{\underline{w}}$ with the Pearson $\underline{r}$. $\kappa_{\underline{w}}$ is equivalent to the product-moment correlation only when partial agreements are weighted in such a way that the agreement weights, $\underline{w}_{\underline{ij}}$'s are according to the formula (derived from Cohen, 1968),

$$\underline{w}_{\underline{ij}} = \frac{(\underline{k} - 1)^2 - \underline{d}^2}{(\underline{k} - 1)^2}$$

where $\underline{k}$ is the number of possible rating categories and $\underline{d}$ is the horizontal or vertical distance from the main diagonal of the $\underline{k} \times \underline{k}$ matrix. The weighting scheme used in the present study typically yields a value of $\kappa_{\underline{w}}$ which is much smaller than the corresponding value of the Pearson $\underline{r}$.

scales independent of scale length (Guilford, 1954). The Scott (1960) homogeneity ratio (HR) is such a measure. The HR, which ranges from 0 to 1, expresses the ratio between two terms. The first term is the difference between the actual variance of the sum of the items and the variance expected if the items were uncorrelated. The second term is the difference between the variance expected if all items were perfectly correlated and the variance expected if the items were uncorrelated. This coefficient can also be interpreted as a weighted average of the scale's interitem correlations.

Convergence

Convergence is the concordance of results of multiple measures between *different* data collection methods. While not explicitly addressed in this definition of convergence, the obverse must also be true: there should be no concordance of results for measures designed to measure different or unrelated concepts. Campbell and Fiske (1959) refer to this supplementary criterion as discriminant validity. Pearson product-moment correlations among observation and interview scales are used to assess the convergence of the observation measures; the Campbell and Fiske (1959) multitrait-multimethod matrix paradigm is used to assess the extent of discrimination among the observation scales.

Results

Repeatability

In the first attempt to develop standardized on-the-job observations, Jenkins et al. (1975) employed a categorization scheme for examining the repeatability of the various items in the observation instrument.

They asserted that when job factors are measured at the same time and at two different times, there are three logical categories into which the results may fall: (1) low agreement between two observers when viewing the same job at the same time and when viewing the same job at different times; (2) low agreement between two observers viewing the same job at different times but high agreement when the job is rated by both observers at the same time; and (3) high agreement between two observers when the ratings are made at the same time and also when they are made at different times. Job factors that are not readily observable (e.g., the job attitudes of employees) should fall into the first category. Job factors that are dynamic (e.g., employee interactions, interruptions) should fall in the second category. Relatively stable job factors (e.g., variety in the job) should fall into the final category. The authors chose the following statistical criteria to categorize their results:

Category 1: κ_w (same time) $< .33$, κ_w (over time) $< .20$

Category 2: κ_w (same time) $> .33$, κ_w (over time) $< .20$

Category 3: κ_w (same time) $> .33$, κ_w (over time) $> .20$

Table 1 contains the results of the repeatability analysis. In the first column are the κ_w's between the two observers viewing the same job at the same time. The second and third columns are the κ_w's between arbitrary pairings of the observation scores for each of the observers in the joint observation with the observation made of the same job/respondent by a single observer. The fourth column is the mean of columns 2 and 3. Column 5 is the κ_w treating all different-time/same-respondent observations as independent measurements.

In general, the measures of repeatability between two observers observing the same job at the same time tend to be lower than those reported in Jenkins et al. Most of the κ_w's for observations made at different

Table 1

Analysis of Observation Items

Scale and Question	Repeatability Estimates[a]					Scale Properties[b]		
	$\kappa_{\underline{w}(\underline{s})}$	$\kappa_{\underline{w}(\underline{d})}$	$\kappa_{\underline{w}(\underline{d})}$	Mean $\kappa_{\underline{w}(\underline{d})}$	$\kappa_{\underline{w}(\underline{d})}$ (combined)	Median $\underline{r}$	Scott H.R.	Coefficient α
Variety						.815	.818	.925
1. How much variety is there in the job?	.421**	.289**	.266**	.278	.277**			
2. The job lets the individual do a variety of different things.	.449**	.363**	.211*	.287	.287**			
3. The job requires that the individual do the same things over and over.[c]	.584**	.324**	.338**	.331	.331**			
Autonomy						.782	.781	.946
4. How much autonomy is there in the job? That is, to what extent does the job permit the individual to decide by himself/herself how to go about doing the work?	.337**	.345**	.262**	.304	.304**			
5. The individual has enough freedom as to how he/she does the work.	.515**	.303**	.373**	.338	.339**			
6. The individual has a lot to say over what happens on his/her job.	.345**	.521**	.530**	.526	.526**			
7. The job allows the individual to make a lot of decisions on his/her own.	.438**	.347**	.370**	.359	.359**			
8. The job denies the individual any chance to use his/her personal initiative or discretion of work.[c]	.204*	.203**	.248**	.226	.225**			
Task Feedback						.733	.733	.844
9. To what extent does doing the job itself provide the individual with information about his/her work performance? That is, does the actual work itself provide clues about how well he/she is doing -- aside from any "feedback" co-workers or supervisors may provide?	.065	.072	-.050	.011	.012			
10. Just doing the work required by the job gives the indivdual many chances to figure out how well he/she is doing.	.100	-.034	-.090	-.062	-.062			

Table 1 continued

Scale and Question	Repeatability Estimates[a] $\kappa_{w(s)}$	$\kappa_{w(d)}$	$\kappa_{w(d)}$	Mean $\kappa_{w(d)}$	$\kappa_{w(d)}$ (combined)	Scale Properties[b] Median *r*	Scott *H.R.*	Coefficient α
Task Impact						.817	.817	.899
11. How much does the work that the individual does on his/her job make a *visible impact* on the materials or objects being worked on or service being rendered?	.219**	-.058	.093	.018	.014			
12. The job allows the individual to make a visible change in the materials worked with or service provided.	.364**	-.065	.097	.016	.016			
Task Completeness						.777	.777	.871
13. How much does the job involve the individual producing an *entire product or an entire service*?	.003	.058	-.028	.015	.015			
14. On the job, the individual produces a whole product or performs a complete service.	.004	-.039	-.122	-.081	-.080			
Task Uncertainty						.642 (.820)[d]	.651 (.834)	.875 (.935)
15. How much *uncertainty* is there in the job?	.503**	.246**	.142	.194	.193**			
16. The job requires the individual to be prepared to handle surprising or unpredictable situations.	.359**	.034	-.020	.007	.007			
17. The job is one that is highly predictable and that rarely presents the individual with surprising or unpredictable situations.[c]	.443**	.216*	.098	.157	.156**			
18. The individual working on the job does tasks which are clearly defined.[c]	.269**	.169	.120	.145	.144*			
Conflicting Demands						.538 (.670)[e]	.546 (.670)	.783 (.802)
19. On the job other people make conflicting demands of the individual.	.296*	.116	-.050	.033	.034			
20. The individual working on this job is free from conflicting demands that others may make of him/her.[c]	.348**	.002	.027	.015	.015			
21. He/She is frequently interrupted for work related reasons.	.305**	-.002	.031	.029	.014			

Table 1 continued

Scale and Question	Repeatability Estimates[a]					Scale Properties[b]		
	$\kappa_{w}(s)$	$\kappa_{w}(d)$	$\kappa_{w}(d)$	Mean $\kappa_{w}(d)$	$\kappa_{w}(d)$ (combined)	Median r	Scott H.R.	Coefficient α
Cooperation						.830	.830	.907
22. To what extent does the job require the individual to *work closely with other people* (either "clients" or people in related jobs within the organization)?	.396**	.184*	.263**	.224	.223**			
23. The individual has to cooperate directly with other people in order to do his/her job.	.460**	.289**	.194	.242	.242**			
Dependence						—[f]	—	—
24. The individual has to depend on the work performed by others in order to get the materials or information he/she needs to do his/her work.	.033	.213*	-.007	.103	.099			
Meaningfulness						—[f]	—	—
25. The job is meaningful	.308**	.270**	.151	.211	.210**			
Required Skills						.722	.771	.905
26. To what extent does the job require the use of *sophisticated* or *complex* skills or knowledge?	.376**	.534**	.414**	.474	.474**			
27. The job requires a high level of skill.	.507**	.552**	.506**	.529	.529**			
28. The job is so simple that virtually anybody could handle it with little or no initial training.[c]	.420**	.506**	.498**	.562	.502**			
Intellectual Demands						.892	.892	.939
29. How *intellectually demanding* is the job?	.411**	.626**	.309**	.468	.467**			
30. The job requires a high level of mental effort.	.512**	.580**	.361**	.471	.469**			
Comfort						.576	.576	.730
31. His/Her work area is clean.	.471**	.289**	.373**	.331	.330**			
32. His/Her job exposes him/her to dangerous or unhealthy conditions.[c]	.289**	.016	-.028	-.006	-.006			

Table 1 continued

Scale and Question	Repeatability Estimates[a]					Scale Properties[b]		
	$\kappa_{\underline{w}}(\underline{s})$	$\kappa_{\underline{w}}(\underline{d})$	$\kappa_{\underline{w}}(\underline{d})$	Mean $\kappa_{\underline{w}}(\underline{d})$	$\kappa_{\underline{w}}(\underline{d})$ (combined)	Median $\underline{r}$	Scott H.R.	Coefficient α
Interruptions						.569	.569	.719
33. He/She is frequently interrupted for work related reasons.	.305**	-.002	.031	.015	.014			
34. He/She is frequently interrupted for non-work related reasons.	.330**	-.070	.172*	.051	.047			
Resource Adequacy						.302	.281	.592
35. He/She is given enough space to do his/her job.	.125	-.156	-.033	-.095	-.092			
36. He/She is given adequate lighting for his/her particular job.	.516**	-.204	-.056	-.130	-.127			
37. He/She has adequate access to machinery, tools or other equipment.	.355**	-.031	.084	.027	.025			
38. The individual working on his/her job frequently had to stop to get things he/she needed and didn't have readily available.[c]	.304**	.126	-.105	.011	-.014			
Physical Effort								
39. The job requires the individual to exert a lot of physical effort.	.367**	.159	.118	.139	.139	—[f]	—	
Dictionary of Occupational Titles						—[g]	—	—
40. Mentoring	-.091	.084	.047	.066	.064			
41. Negotiating	.445**	.200	.317*	.258	.261**			
42. Instructing	.636**	.203	.443**	.323	.330**			
43. Supervising	.455**	.504**	.416**	.460	.457**			
44. Persuading	.141	.261*	.283*	.272	.274**			
45. Speaking-Signaling	.396**	.365**	.402**	.384	.382**			
46. Serving	.187	.164	.012	.088	.091			

Table 1 continued

Scale and Question	Repeatability Estimates[a] $\kappa_{w(s)}$	$\kappa_{w(d)}$	$\kappa_{w(d)}$	Mean $\kappa_{w(d)}$	$\kappa_{w(d)}$ (combined)	Scale Properties[b] Median r	Scott H.R.	Coefficient α
Interactions						—[g]	—	—
How often does the individual interact verbally with . . .								
47. His/Her supervisors.	.494**	.329**	.128	.229	.228**			
48. His/Her co-workers (at the same level of the organization as the individual).	.355**	.220*	-.018	.101	.102			
49. His/Her co-workers (at different levels of the organization from the individual).	.449**	.235**	.212*	.224	.224**			
50. Others (customers, clients, patients).	.184*	.261**	.016	.139	.137			
51. Unidentifiable individuals (can't tell who they are).	.148	.045	.000	.023	.023			

NOTES:

[a] $\kappa_{w(s)}$ is the weighted kappa statistic of agreement between two observers observing the same job/respondent at the same time. These values are based on a total N of 44 with exact N's varying from 40 to 44. $\kappa_{w(d)_1}$ and $\kappa_{w(d)_2}$ represent agreement between two observers observing the same job/respondent at different times. Pairing of observers was arbitrary. Both statistics are based on a total N of 41 with exact N's varying from 38 to 41. Mean $\kappa_{w(d)}$ is the average of $\kappa_{w(d)_1}$ and $\kappa_{w(d)_2}$. $\kappa_{w(d)}$ combined treats all observations of the same job respondent at differnt times independently and is based on a total N of 82 with exact N's varying from 76 to 82.

[b] Median r, Scott's homogeniety ratio and Cronbach's alpha are based on the relationships of the coexistent items of the scales over the total sample of jobs/respondents observed ($N = 147$).

[c] Reversed scoring

[d] Statistics when variable 16 is eliminated from the scale.

[e] Statistics when variable 21 is eliminated from the scale.

[f] Single item scale - statistics not possible to compute.

[g] Scale not constructed to be homogeneous - statistics not computed.

*$p < .05$ (two-tailed test)

**$p < .01$ (two-tailed test)

times approximate the magnitude of those reported for the same items in the previous paper, although some are higher and some lower. If the categorization scheme described above is applied to these results, of the 31 items included in both instruments, only seven are categorized differently in this phase compared to the first phase. The previous results are strongly replicated ($X^2 = 24.180$, df = 4, N=31) [see Table 2].

Of the new items and scales added to this phase of the research, the operations of task impact, task completeness, interruption by others and unidentifiable individuals, and the DOT classification of mentoring, persuading and serving were not repeatable and fell into category three. The remaining new operations were repeatable, either when observations were made at the same time or both when made at the same and different times.

Homogeneity

Table 1 also contains, when appropriate or possible to compute, the median inter-item correlation, Scott's homogeneity ratio and coefficient α for each a priori scale. Of the multi-item scales designed to be homogeneous, only conflicting demands, comfort, interruptions and resources failed to meet the .60 criterion used in the first phase of this study. When one item is removed from the conflicting demands scale, it surpasses this standard (HR = .670) and the comfort scale approaches the criterion level.

Convergence

Eight of the scales can be reasonably examined for convergence since they were measured in both the interview and by observations and exhibited

Table 2

Percentage Agreement of Empirical Item Classification between 1973 and 1975 Observation Investigations

		1975 Investigation Category		
		I	II	III
1973 Investigation	Category I	32.3%	3.2%	6.5%
	II	0.0%	19.4%	6.5%
	III	3.2%	6.5%	22.6%

N = 31, *df* = 4

$\chi^2 = 24.188$ ($p < .0001$)

repeatability of their constituent items. Table 3 presents the multitrait multimethod matrix for the eight scales. Entries on the main diagonal of the matrix are internal consistency reliability estimates for scales within a given method. Entries marked by asterisks were single item scales for which internal consistency reliabilities are impossible to compute. The correlations in parentheses in the lower left of the matrix represent the convergence between the same construct measured by two methods. Five of the eight convergence correlations are significant at the .01 level: variety, autonomy, required cooperation, required skill and intellectually demanding. Two of these, autonomy and required cooperation, were significantly higher ($p < .05$) than in the previous study. None was significantly lower. Uncertainty, conflicting demands, and physically demanding failed to demonstrate adequate levels of convergence. It appears that what the observers saw as uncertainty and conflicting demands, the respondents perceived as required cooperation, required skill and intellectual demand. What the respondents viewed as variety and autonomy, the observers scored as uncertainty and conflicting demands. When respondents reported their jobs as being physically demanding, observers tended to report those jobs to be lower in autonomy, skill required, uncertainty, and intellectual demand.

With respect to the discrimination among constructs, the median intercorrelation among constructs in the observation monomethod triangle decreased from .71 in the first study to .53 in the present study (c.f., Jenkins et al., 1975) With respect to the other criteria of discriminant validity, the interview and observation methods both fail to meet Campbell and Fiske's criteria.

Table 3

Multitrait-Multimethod Matrix of Convergence and Discriminant Correlations among Scales

	Observation								Interview							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Observation:																
1. Variety	.925															
2. Autonomy	.691	.946														
3. Uncertainty	.623	.680	.935													
4. Conflicting demands	.336	.354	.525	.802												
5. Required cooperation	.559	.453	.646	.435	.907											
6. Required skill	.568	.693	.740	.360	.570	.905										
7. Intellectually demanding	.625	.763	.776	.447	.599	.896	.939									
8. Physically demanding	.052	-.072	.045	-.101	-.046	-.055	-.067	*[a]								
Interview:																
1. Variety	(.285)	.286	.305	.306	.344	.269	.322	-.092	.609							
2. Autonomy	.365	(.512)	.451	.419	.403	.333	.427	.004	.441	.860						
3. Uncertainty	.018	-.077	(-.070)	-.161	-.043	-.005	-.018	-.185	.057	-.269	.576					
4. Conflicting demands	.030	.038	.181	(.055)	.214	.208	.185	-.051	.022	-.137	.267	.333				
5. Required cooperation	.276	.377	.334	.341	(.368)	.362	.413	.062	.260	.345	-.130	.041	*[a]			
6. Required skill	.322	.410	.317	.248	.354	(.443)	.453	-.011	.413	.406	-.178	.050	.309	.760		
7. Intelluctually demanding	.311	.460	.426	.315	.399	.439	(.433)	.002	.361	.522	-.122	.200	.356	.648	.714	
8. Physically demanding	-.118	-.243	-.232	-.151	-.120	-.251	-.318	(.149)	-.004	-.152	-.115	.183	-.147	-.100	-.015	*[a]

N = 147

[a] Single item scales -- internal consistency reliabilities are not possible to compute.

Discussion

This investigation was an effort to replicate and refine the earlier study with respect to standardized on-the-job observations as an alternative technique for data collection. The refinement of the method was intended to increase the convergence of the constructs with interview measures of the same construct and to increase the discrimination among constructs when measured by observations, without sacrificing either the stability or homogeneity of the measures. The attempt was in large measure successful.

First, there was a substantial improvement (reduction) in the extent to which different constructs measured by observations were interrelated. Others have similarly found that job characteristics that are theoretically independent are not independent in their value distributions among "real world" jobs (Hackman & Lawler, 1971; Hackman & Oldham, 1975, 1976; Jenkins et al., 1975). While some part of such obtained covariance may be due to methods' similarity or observer halo effects, it is likely that much of it is valid ecological correlation, i.e., arising from non-random, clustered distribution of the job attributes within a population of jobs. Such ecological correlations were also present among constructs as measured by the interviewer though they were not as substantial. The improvement in discrimination undoubtedly resulted from the concepts provided the observers. In addition, because observers were selected in part on the extent to which their ratings of sample jobs agreed with expert ratings of those jobs, the selection techniques probably added to the extent that the observers were able to distinguish between, for example, variety and autonomy or conflicting demands. This increased discrimination among constructs was achieved without an important loss with respect to the extent that the constructs were measured with stability (i.e., they were repeatable) and homogeneity.

The verdict with respect to convergent and discriminant validation would have been stronger, however, had the three Campbell and Fiske criteria of discriminant validity been met more fully. First, convergence correlations frequently failed to exceed correlations in the same row or column of the monomethod triangles. Second, the convergence correlations failed to exceed elements in the same row or column of the heteromethod triangles. Thus, applying the strict standards of Campbell and Fiske, progress does seem to have been made with respect to reducing the intercorrelations among the constructs when measured by observation and generally increasing the values of the convergence diagonal, but much still remains to be done with respect to discriminant validity.

With respect to repeatability and homogeneity, the modifications in training, selection, and instrument design did reduce the values of the $\kappa_{\underline{w}}$'s and the homogeneity ratios somewhat. For the most part, however, these decreases were not significant, and substantially replicated the findings of the previous study as to data repeatability and homogeneity.

In sum, this attempt to refine the observation methodology has been reasonably successful, with certain distinct gains. The method is far from being fully satisfactory, however, and several problems remain to be resolved.

Interactions of Job and Job Holder/Observer

While it appears possible, with the use of concepts training and practice, to train observers to recognize some task dimensions such as task feedback, task impact or task completeness and report measures of the concepts consistently (i.e, the measures are internally consistent across observers), observers do not develop a shared metric with respect to these concepts. Observers fail to agree that there is a given level of these concept-attributes present in a particular job even when they view the job at the same time. It may be that such concepts are not inherent dimensions of a job, but rather the result of a complex interaction of other task characteristics and the job holder. When observers attempt to report levels of these characteristics they must then inject their own personalities, needs, and desires into the observations much as job holders must do when they report the perceived amount of these concepts they experience in their jobs. Such an interaction would inevitably lead to a lack of standardization of measurement and consequently reduce the degree to which the measures of the concepts could be stable or repeatable, even though, within a given observer, the measures of the concept are consistent.

Differential Observability of Jobs.

In the previous study, it was asserted that some jobs were more inherently observable along certain dimensions than others. Because of a lack of appropriate data and insufficient cases it was not possible to test this assertion. The present investigation still suffers from small numbers of cases in different job categories, but has more of the classification information required to begin examining this assertion. Such data include codes from the Dictionary of Occupational Titles, Census Occupational Classifications, and supervisory status. While we still believe that some jobs lend themselves more readily to valid job observations, examination of the data when classified along the above dimensions failed to yield any interpretable or confirming results. Our search for a categorization scheme to distinguish observable from non-observable jobs on either empirical or conceptual basis will continue.

Implicit Theories of Job Structure

The correlations in the heteromethod triangles remain high and seriously reduce the discriminant validity of the observation method. These correlations may be artifactual, the product of the high correlations in the observation method triangle and the often sizeable correlations in the interview method triangle and the convergence diagonal. A second explanation, though untestable in retrospect, is also possible. There is increasing evidence that individuals have implicit theories about the relationships that exist among organizational constructs (Staw, 1975; Eden & Leviatan, 1975). These implicit theories constitute a set of presuppositions regarding relationships among constructs that observers bring with them to the observation situation. In the absence of interpretable cues, observers apply their implicit theories to the job stimuli and rate jobs they observe to be high on one job dimension as being high on other dimensions as well. For example, it appears that when observers rate a job as being high on autonomy and variety, they also rate it as being higher along the other dimensions (except physically demanding). Using the notion of implicit theories, it may be argued that when observers see variety and autonomy in a job, they imbue that job with increased uncertainty and conflicting demands. Since both uncertainty and conflicting demands show low convergence between methods, we must assume that observers have difficulty observing them and, for want of interpretable cues, use their implicit theories to guide their ratings. Explicit attempts to disavow the observers of these potential implicit theories were not made during the observation training. It is, therefore, highly likely that observers carried these implicit theories to the observation setting and employed them when rating ambiguous stimuli. Efforts should be made in the future training of observers to disabuse them of these theories. Furthermore, observers should be encouraged not to infer ratings on one job characteristic from their ratings of another job characteristic.

Job Redefinition Process

While the modification in the training procedures appears to have the effect of increasing some of the values on the convergence diagonal (Table 3), they remain much lower than what might normally be expected and desired. Why were they not higher? One explanation may be found in a process hypothesized by Hackman (1969): job redefinition. He proposed that when a task (or in this case, job) is given to an individual, the performer routinely "redefines" the task to be consistent with his or her needs, goals, and values before performing it. This redefinition may be cognitive; that is, the objective characteristics of the job may not be changed but merely distorted by the job holder so that he "experiences" more of the particular characteristics he wants to experience. In other cases, the job holder may actually introduce more of a given characteristic to a job and introduce it in such a way that it would not be apparent to an observer. Reports are legion of workers adding unrequired variety to their jobs or gaining some autonomy by fashioning production techniques different from those prescribed. From a different theoretical perspective, Alderfer (1972) proposes that a performer who desires or "needs" more of a given task characteristic may be more adept at exploiting a given level of that characteristic in a job and hence experiences more of it. Thus, in all cases, the job as experienced by the job holder is different from the job the observer sees and a reduction in the convergence between reports of the job as perceived by the job holder and the observer is a consequence. The use of an individual difference moderator such as high order need strength or importance ratings of task characteristics, may serve to increase the convergence between the two methods of measurement.

Future Role of Job Observations

Currently the most plausible model of the relationship between the characteristics of jobs and employee behavior is that of Hackman and Oldham (1975, 1976). They posit that the perceived characteristics of jobs create "critical psychological states" for the individual job holder which in turn lead to individualistic employee behavior and affective reactions. Presumably, though not explicitly included in their model, objective job characteristics interact with the desires, needs, and goals of the individual job holder to produce experienced job characteristics. Since it is the objective characteristics of jobs which are manipulated in job redesign efforts, the standardized job observation methodology has an important place in diagnosing areas for job redesign, evaluating the magnitude and nature of the redesign efforts, and understanding the effects of the nature of jobs on the individuals who perform them. As in the case of other research issues involving both persons and their environments, it appears to be essential to treat both objective and subjective variables jointly.

References

Alderfer, C. P. *Existence, relatedness, and growth: Human needs in organizational settings*. New York: The Free Press, 1972.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Cohen, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 1960, *20*, 37-46.

Cohen, J. Weighted kappa: Nominal scale agreement with provision for scale disagreement or partial credit. *Psychological Bulletin*, 1968, *70*, 213-220.

Eden, D., & Leviatan, U. Implicit leadership theory vs. a determinant of the factor structure indulging supervisory behavior scales. *Journal of Applied Psychology*, 1975, *60*, 736-741.

Fleiss, J. L., Cohen, J., & Everitt, B. S. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 1969, *72*, 323-327.

Guilford, J. R. *Psychometric Methods*, 2nd edition. New York: McGraw-Hill, 1954.

Hackman, J. R. Toward understanding of the role of tasks in behavioral research. *Acta Psychologica*, 1969, *31*, 97-128.

Hackman, J. R., & Lawler, E. E. III. Employee reactions to job characteristics. *Journal of Applied Psychology*, 1971, *55*, 259-286.

Hackman, J. R., & Oldham, G. R. Development of the job diagnostic survey. *Journal of Applied Psychology*, 1975, *60*, 159-170.

Hackman, J. R., & Oldham, G. R. Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, 1976, *16*, 250-279.

Jenkins, G. D., Jr., Nadler, D. A., Lawler, E. E. III, & Cammann, C. Standardized observations: An approach to measuring the nature of jobs. *Journal of Applied Psychology*, 1975, *60*, 171-181. Also Chapter 4.

Kerlinger, F. N. *Foundations of behavioral research*. New York: Holt, Rinehart, and Winston, 1964.

Scott, W. A. Measures of test homogeneity. *Educational and Psychological Measurement*, 1960, *20*, 751-757.

Staw, B. M. Attribution of the "causes" of performance: A general alternative interpretation of cross-sectional research on organizations. *Organizational Behavior and Human Performance*, 1975, *13*, 414-432.

References (continued)

U. S. Bureau of the Census. 1970 census of population: Alphabetical index of industries and occupations. Washington, D. C.: U. S. Government Printing Office, 1971.

U. S. Department of Labor. Dictionary of occupational titles. Washington, D. C.: U. S. Government Printing Office, 1965.

Chapter 8

IDENTIFYING SOURCES OF INSTABILITY IN MEASURES OF WORKING CONDITIONS AND WORK-RELATED ATTITUDES AND BEHAVIORS

by

Bernard Goitein

ABSTRACT

Organizational studies typically encounter difficulties in arranging a random assignment of subjects to experimental conditions; hence the best quasi-experimental design available is often the non-equivalent control group (NECG) design. The error variance term of such a design is the variance of a change score. Kenny (1975) proposes four types of change score for NECG designs. The present analysis computes the variances of these four change scores for 17 measures of working conditions, worker attitudes and behaviors. It uses 272 cases followed over a lag of approximately 20 months.

Reliabilities of these change scores are reported and it is found that under "control group" conditions, relatively little reliable change is observed in measures of working conditions, whereas fairly high reliabilities are found for worker attitude and behavior change scores.

Correlations are computed and methodological guidelines presented for the selection of potentially valuable covariates from among the intercorrelated variables. Such covariates may be used in future studies to reduce the error variance term of one or another of the 17 work-related measures.

Some particular types of jobs, job conditions and worker subpopulations are specified that tend to display more or less variance of change scores and hence are likely to generate larger or smaller error variance terms of work-related measures.

Also studied is the stability over time of structures of intercorrelations of work-related variables. The structure of relations as depicted by Smallest Space Analyses was found to be quite stable over the investigated lag. Certain results from the earlier analyses are then considered in terms of the relative importance of the person (worker) and the situation (job) in determining work-related experiences. The particular occupant of a job was found to have much greater impact on scores than the particular job of a worker.

Chapter

IDENTIFYING SOURCES OF INSTABILITY IN MEASURES OF WORKING CONDITIONS AND WORK-RELATED ATTITUDES AND BEHAVIORS

Introduction

Experiments in job change appear to be increasing in number and are likely to continue to do so in the foreseeable future. They will probably be most common with regard to working hours, job content, and the multitude of changes subsumed under the rubric of "organizational development."

To date, many of the most widely cited experiments with working hours and job content have failed to meet even the most rudimentary criteria of adequate experimental design. Rigorous sampling of situations and random assignment of subjects are rare, as are adequate control groups. If, however, one refrains from conducting any kind of field experiment in employing establishments until the ideal experimental opportunity comes along, very few experiments with working conditions will ever get done.

Campbell and Stanley (1963) distinguish twelve sources of invalidity in experiments. These, together with descriptions of six prototypical experimental and "pre-experimental" designs, are reproduced in Table 1. The symbols describing each design are as follows: R=randomization of subjects; X=administration of the experimental manipulation; O=administration of measurement. There is, unfortunately, a plethora of "pre-experimental" designs among studies of job change. Most of the studies

Table 1

Sources of Invalidity in Pre-experimental and Experimental Designs

	Sources of invalidity											
Type of design	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection and maturation, etc.	Interaction of testing and X	Interaction of selection and X	Reactive arrangements	Multiple-X interference
Pre-Experimental designs:												
1. One-Shot Case Study X 0	-	-				-	-			-		
2. One-Group Pretest-Posttest Design 0 X 0	-	-	-	-	?	+	+	-	-	-	?	
3. Static-Group Comparison X 0 0	+	?	+	+	+	-	-	-		-		
True experimental designs:												
4. Pretest-Posttest Control Group Design R 0 X 0 R 0 0	+	+	+	+	+	+	+	+	-	?	?	

Table 1 - (Cont'd)

Sources of Invalidity in Pre-experimental and Experimental Designs

Type of design	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection and maturation, etc.	Interaction of testing and X	Interaction of selection and X	Reactive arrangements	Multiple-X interference
	Sources of invalidity											
5. Solomon Four-Group Design R O X O R O O R X O R O	+	+	+	+	+	+	+	+	+	?	?	
6. Posttest-Only Control Group Design R X O R O	+	+	+	+	+	+	+	+	+	?	?	

NOTE: A minus indicates a definite weakness, a plus indicates that the factor is controlled, a question mark indicates a possible source of concern, and a blank indicates that the factor is not relevant.

SOURCE: Campbell and Stanley, 1963.

on job enrichment and/or enlargement reported in the Work in America (1973) appendix are usually of types 1 and 2, occasionally 3, and a few used no systematic measurement at all. Glickman and Brown's (1973) review of experiments in working hours shows an abundance of type 3 designs, after-the-fact assessments of workers who did and did not volunteer to work under the new time arrangements. A major obstacle to applying "true" experimental designs (designs 4-6 in Table 1) in studies of employing establishments is the difficulty of randomization. While on a group or unit level one might, for example, randomly assign branch offices of a firm (or work groups within the firm) to experimental and non-experimental conditions; this is less feasible on an individual level. It is impossible where, for example, the "groups" exposed to the conditions are two plants of the same firm and individuals have selected themselves (or been selected) into the two plants.

In the latter case, a not uncommon one in studies of employing establishments, Campbell and Stanley suggest one commonly used quasi-experimental design, the "non-equivalent control group design." It does not assume random assignment of people, and its form, in the language of Table 1, is:

O x O

O O

The design, according to Campbell and Stanley, is applicable where "the groups constitute naturally assembled collectivities such as classrooms, as similar as availability permits but yet not so similar that one can dispense with the pretest" (p.47). Statistical evaluations of the experimental treatment are made not in terms of experimental and control groups but in terms of the relative gain of the experimental group--that is, the amount of improvement in the experimental group above and beyond that which has occurred "naturally" in the control group during the same period.

In spite of the limitations of this design, it seems an unrealistic purism to postpone experiments in work organizations because their designs cannot conform to the standards of the "true" experimental designs appearing in the lower half of Table 1. Given less than infinite resources, problems of access to employing establishments and local constraints existing at these sites, the non-equivalent control group design is likely to be around for quite some time. If the majority of the job change studies reported in the Work in America appendix had used designs even as good as this one, their "findings" would have been more conclusive.

We believe, however, that the non-equivalent control group design can be used more efficiently in future experiments in work organizations than it has in the past. An efficient design is one which obtains the greatest information for the least cost. Critical to the cost of experiments in organizations is sample size. The sample sizes must be sufficiently large to support statistically a failure to disconfirm the null hypothesis lest it be concluded without justification that an experimental treatment has had no effect (i.e., lest a Type II statistical error be made). Given a certain size of effect, a standard answer to the question of sample size and statistical power is "the bigger the sample the better." Beyond a certain point, however, increments in sample size become inefficient. Increases in precision beyond this point simply do not justify the additional expense entailed in enlarging the sample. For example, standard errors of estimate in national probability household samples in excess of about 1500 decrease at a rate that does not always justify the expense required to add more respondents (except, of course, where greater precision is wanted for subpopulations within the total population). Samples become inefficient when they have more than enough cases to test reliably a study's major hypotheses. It is like using a hammer to swat a fly.

In order to determine efficient sample size for experiments in work organizations using non-equivalent control group designs, an advance estimate of what is likely to happen to the control group is essential. At present this knowledge remains elusive. Though control groups have been used in countless experiments, there has been no accumulated knowledge of what constitutes the "natural" events that commonly occur in modern American worksettings or of how and why measures of work-related attitudes and behaviors change over time even when untouched by experimentation.

Analysis Goals

The present study will attempt to provide some of this needed knowledge about control groups. The specific goals of the proposed analyses are:

1. to identify the amount of instability that occurs over time in measures of working conditions, workers' attitudes, and workers' behaviors among a sample of workers whose jobs have been subject to no experimental change;
2. to identify some of the sources contributing to this instability;
3. to identify situations where the measures are most and least stable.

These analyses will provide future experimenters with some indication of how much the measures vary over time. They will also identify those situations where at least some measures are comparatively stable so the experimenter might economize on the size of his or her control group. Where, on the other hand, "natural" change is rampant, and the proposed experimental change is not expected to have overwhelming effects, several alternatives are possible. One is simply to increase the sample size. Another would be to reduce the variance resulting from measurement error. Estimates of the variance attributable to errors of measurement will be presented for most of the measures used. A third alternative would be to use statistical control techniques (e.g., the use of covariates) to reduce the variance term by removing what is, from the experimenter's point of view, error variance. Potentially valuable covariates will be identified in this study. And, finally, if the cost of these activities (increasing sample size, improving reliability, and the use of suitable covariates) does not justify the value of the experiment, the study should simply be scrapped. The availability of estimates of the amount of "natural" change to be expected, and the experimenter's assessment of the strength of the manipulations will allow the decision on how to proceed to be made using specific information, rather than relying on the experimenter's guesswork alone.

Method

Sample

The overall design of the two phase study of which this investigation is a part, is described in Chapters 1 and 2. The present report is based upon data from all the 272 cases who participated in the two phases of the study. These 272 cases were interviewed in three of the five midwest organizations originally studied (two automobile parts manufacturers and a hospital), about 20 months after the first phase of the study.

The 272 cases may be subdivided into three groups:

1. 163 cases where a worker was interviewed for Phase II who had been interviewed for Phase I, and who was still working in the same job he or she had had in Phase I;

2. 73 cases where a worker was reinterviewed for Phase II, but who had changed jobs in the period between Phase I and Phase II;

3. 36 cases where a worker had not been interviewed at Phase I, but who was interviewed for Phase II because he or she was filling a job vacated by a Phase I respondent.

The classification of jobs as different or the same was performed by coders familiar with the research sites, on the basis of information from company records (job title, work group, department and shift) on the jobs at each phase.

Ideally, the identification of "normal" change in working conditions, workers' attitudes, and workers' behaviors would be based upon data obtained from a national probability sample of workers—such as those available from the 1972-73 Quality of Employment Survey (Quinn and Shepard, 1974). The latter survey, however, lacks follow-up over time, a central component of the analysis proposed here.

The demographic distributions (on age, sex, race, education and occupational group) of the 272 Phase I respondents who were to be reinterviewed (or whose job replacement was to be interviewed) are, fortunately, fairly similar to the demographic distributions found for the national sample of workers (see Table 2), though blacks, women, and those with less education are somewhat over-represented in the present sample.

Measures

The stability of measures relevant to studies of the effects of working conditions will be assessed with regard to three types of variables: workers' perceptions of working conditions, worker attitudes, and worker behavior. Two data sources are used: worker interview responses and supervisory ratings. Other sources of information (on-the-job observations and company records) were not used since the data were either not comparable over the two phases, not comparable over the sites, or (for one variable) judged unreliable.

With one exception discussed below, all measures in the three categories were created from the pool of items common to both phases of the study, in order to ensure comparability of measurements between phases. Wherever possible, measures were scored in the same direction so that "high" meant "good" working conditions and "positive" work-related attitudes and behaviors (e.g., high satisfaction, high involvement, low ambiguity of work role, and few days absent).

Working Conditions. Working conditions were measured by the Quality of Employment (QoE) Facets' scales (Comfort, Financial Rewards, Resource Adequacy, and Challenge--Barnowe, Mangione and Quinn, 1972), and the overall QoE/Total measure (Cobb, 1973). The scales from the 31 items used in the present study are described in greater detail in Chapter 3 of this report.

A four item measure of role ambiguity devised by Beehr (1976) was also used.

Worker Attitudes. Various aspects of worker attitudes were measured.

Table 2

Demographic and Occupational Distribution of Effectiveness in Work Roles Cross-Phase Subsample and Comparison National Survey Sample

Demographic or occupational characteristic	Percentage distribution	
	Effectiveness in Work Roles sample (N=272)	Comparison national sample[a] (N=2157)
Sex		
Men	56.3%	62.1%
Women	43.8	37.9
Age		
Under 21	9.2	8.1
22-29	28.7	27.1
30-44	32.7	30.6
45-54	19.9	20.6
55 or older	9.6	13.6
Education		
Some grade school or less	3.7	4.7
Completed grade school	8.9	6.6
Some high school	21.4	14.2
Completed high school	33.9	38.4
Some college	22.1	20.9
College degree or more	10.0	15.2
Race[b]		
White	76.3	91.5
Black	23.7	8.5

Table 2 (continued)

Demographic and Occupational Distribution of Effectiveness in Work Roles Cross-Phase Subsample and Comparison National Survey Sample

Demographic or occupational characteristic	Percentage distribution	
	Effectiveness in Work roles sample (N=272)	Comparison national sample (N=2157)
Major occupation group[c]		
Professional and technical	18.4%	14.3%
Managers, officials, and proprietors	9.2	9.8
Clerical	12.1	15.1
Sales	0.0	0.2
Craft workers and foremen	9.6	15.8
Operatives	30.9	29.0
Service workers, excluding private household	16.2	12.6
Laborers	3.7	3.1

---

a Quinn and Shepard (1974).

b Excludes minority races other than blacks.

c Based on 1960 Census codes. The Effectiveness in Work Roles sample contained no farmers or private household workers, and workers in these occupations have been excluded from the bases of the comparison national statistics.

Measures of Job Satisfaction (JS) facets first developed in the 1969-70 Survey of Working Conditions (Quinn, Seashore, Kahn, Mangione, Campbell, Staines and McCullough, 1971), were used here. These facets parallel the QoE facets, with one added facet (coworkers), and therefore include: Satisfaction with Comfort, Satisfaction with Financial Rewards, Satisfaction with Resource Adequacy, Satisfaction with Challenge, and Satisfaction with Coworkers (henceforth abbreviated as JS/Comfort, JS/Financial Rewards, JS/Resource Adequacy, JS/Challenge, and JS/Coworkers).

The items used to measure some of these job satisfaction facets differ slightly from those used in national survey work (Quinn, et al., 1971; Quinn and Shepard, 1974), because of the limited item pool of the present study. While JS/Comfort, JS/Financial Rewards and JS/Coworkers are measured as described in Quinn, et al. (1971) and JS/Challenge is measured as described in Quinn and Shepard (1974), JS/Resource Adequacy consists of only three of the four items originally used by Quinn et al. (1971). The three items measure satisfaction with amount of information, equipment, and supervisory competence, while the omitted fourth item measures satisfaction with the clarity of the definition of worker responsibilities.

Other attitudinal measures used include a global, facet-free job satisfaction index (a five item measure of satisfaction with the job as a whole, developed and used in national survey work, Quinn and Shepard, 1974), and a measure of job involvement. Among the pool of items common to both phases, only one item on job involvement (previously used by Beehr, 1974, and based upon an item developed by Pelz and Andrews, 1966) was available:

> Some people are completely involved in their job--they are absorbed in it night and day. For other people, their job is

simply one of several interests. How involved do you feel in your job--very little, slightly, moderately, or strongly involved?

This one item is the measure of involvement used in both phases for most analyses. Since some of these analyses permit non-identical measures of the same construct and since a multi-item measure is typically more reliable than a one-item measure, an index of involvement was developed that included items from outside the pool of items common to both phases, i.e., items unique to one phase of the data collection. Two such items found in the Phase II pool of items were unique to that phase. A three item index of job involvement was created for Phase II, then, along with the aforementioned single item used in both phases. This three item measure consists of the one item measure plus two additional items with responses on a four point scale form "very true" to "not at all true":

- I live, breathe, and eat my job.
- The most important things that happen to me involve my job.

Cronbach's alpha was computed for this three item scale, and found to be 0.71.

Work-related behavior. Various aspects of work-related behavior were likewise measured. Beehr's (1974) two item measure of effort-expended (referred to as "motivation to work" by Beehr) was used. The two items read:

- Would you say you work harder, less hard, or about the same as other people doing your type of work?
- How often do you do some extra work for your job which isn't required of you? Would you say you do this often, sometimes, rarely, or never?

Single item measures of reported number of days absent in the month preceding the interview (henceforth referred to as absences), and reported likelihood of looking for a new job with another employer in the next year (henceforth referred to as intent to turn over) were also used. Supervisory ratings of workers on eight aspects of their work behavior (e.g., dependability, punctuality, quantity and quality of work) were combined to yield an overall supervisory rating of worker quality (this variable is henceforth referred to as supervisory rating).

Results and Discussion

The present analysis of instability of measures over time is divided into four parts, the first three corresponding to the three major analysis goals cited in the introduction, and a fourth dealing with the stability of the structure of interrelations over time.

1. Quantifying the Amount of Instability Over Time

In order to quantify the amount of variance of scores over time (i.e., the error variance for an experimental study), the problem of what score to use to compute that variance must be confronted. The problem of the most suitable way to score change has been the object of a long and complex controversy that extends over the last fifteen years (e.g., Cronbach and Furby, 1970; Harris, 1963; Kenny, 1975; and Werts and Linn, 1970a). A recent review (Kenny, 1975) deals explicitly with the problems of the computation of change scores for the non-equivalent control group design. Kenny's paper is particularly useful here, as the users of the non-equivalent control group design are a major target audience for the present paper. Kenny (1975) concludes that rather than one best way to measure change, one of four suggested solutions to the problem should be used, depending on the problems of the particular experimental or quasi-experimental design. To assist future users of non-equivalent control group designs, the variance of each of the four change score solutions will be reported for each variable as estimates of that variable's variance over time. The four measurement techniques proposed are:

1. Raw gain (time two minus time one score);

2. Gain of standardized scores, i.e., scores are standardized (z transformation) separately in each phase, and time one is then subtracted from time two;

3. Time two scores residualized on time one (essentially equivalent to the dependent variable used in analysis of covariance, where the time one score is the covariate and the time two score is the dependent variable);

4. Time two scores residualized on time one, correcting for the unreliability of the time one score (equivalent to the analysis of covariance solution in 3 [above], but where the regression coefficient is divided by the reliability of the time one measure). While Kenny derives this measure from an earlier discussion by Lord concerning the analysis of covariance, it is identical to the independent true gain measure proposed by Tucker, Damarin and Messick (1966).

The variances over time were computed for each of the four change scores (raw gain, gain of standardized scores, residual, and corrected residual) for each of the 17 measures of working conditions, work-related attitudes and behavior. The change scores themselves were not generated for these computations. Rather, the variance estimates were computed from the appropriate within-phase variances and across-phases covariances. The derivations of the formulae used may be found in Appendix 1. The variance estimates and their components appear in Table 3. With reliability estimates (derived in the following section) available for only 14 of the 17 measures, variances of the residual scores corrected for unreliability were not generated for the one item measures of involvement, absences, and intent to turnover. Involvement was measured in this analysis with the same single item measure in both phases. The usefulness of these variance estimates is limited largely to potential users of these particular measures. Insofar as some of these measures are relatively standard and have been repeatedly used in the past (e.g., global satisfaction, or the QoE facets), these estimates are likely to be useful. Some evidence regarding the generalizability of these variances is available from Quinn and Shepard (1974), whose national survey data were collected a few months after Phase I of the present study. Quinn and Shepard report the standard deviations of three variables that are also studied here. The reported standard deviations of facet-free job satisfaction and intent to turnover (pp.80-81) are very close to those found in Phase I of the present study (Table 3) and the standard deviation of QoE/Total is only slightly greater here (0.49) than in the national sample (0.44, pp.245-246).

Certain overall observations concerning these results may be made. The values of the residual gain scores are always the lowest variance of any form of change score, as expected by the least squares procedure used in designing linear regression. The variances of the corrected residual scores are typically close to the values of the residual variance estimates,

Table 3

Variance of Change Scores and Components of Variance of Change

Variable	N	Variance at time one	Variance at time two	Time one-time two correlations	Reliability at Phase I	Variance of raw gain	Variance of standardized gain	Variance of residual gain	Variance of corrected residual gain
QoE/Resource adequacy	252	0.51	0.60	.32	.60	0.75	1.35	0.54	0.56
QoE/Financial rewards	252	0.46	0.56	.50	.60	0.51	0.99	0.42	0.48
QoE/Challenge	260	0.68	0.70	.59	.60	0.56	0.81	0.46	0.57
QoE/Comfort	256	0.30	0.32	.46	.60	0.33	1.08	0.25	0.28
QoE/Total	259	0.24	0.30	.60	.66	0.22	0.79	0.19	0.22
Ambiguity	260	0.36	0.35	.30	.54	0.50	1.39	0.32	0.34
Global job satisfaction	268	0.99	1.00	.46	.73	1.06	1.07	0.78	0.81
JS/Coworkers	268	0.38	0.33	.25	.49	0.54	1.50	0.31	0.33
JS/Challenge	267	0.44	0.45	.54	.83	0.40	0.91	0.32	0.32
JS/Comfort	267	0.27	0.30	.48	.66	0.30	1.04	0.23	0.25
JS/Resource adequacy	268	0.44	0.45	.46	.66	0.39	1.08	0.36	0.39
JS/Financial rewards	267	0.46	0.51	.44	.66	0.54	1.12	0.41	0.44
Effort expended	260	0.56	0.53	.45	.44	0.61	1.11	0.43	0.60
Supervisor rating	60	1.03	1.56	.43	.92	1.51	1.15	1.28	1.28
Absences	268	12.94	8.01	.13	--[a]	18.21	1.73	7.86	--[b]
Intent to turnover	267	1.69	1.96	.37	--[a]	2.30	1.26	1.69	--[b]
Involvement	264	0.77	0.74	.41	--[a]	0.89	1.17	0.62	--[b]

[a] No reliability estimate available.

[b] Variances of corrected residual gain cannot be computed without reliability estimates, which are not available for these variables.

except for the effort-expended scale whose reliability in Phase I (0.44) is low. The corrected residual score variance of the effort-expended scale is, in fact, larger than the variance of the time two scores of this variable, so variance seems to have been added to these scores, rather than removed, because of the low reliability of the measure. It is easily shown that this occurs whenever a measure's reliability is less than 0.50.

The variance of raw gain is always particularly large when the correlation of a variable with itself over time is low. This phenomenon is particularly marked in the case of the absences variable, which correlates with itself over time only 0.13. This measure would seem to be quite unstable, with substantial natural change to be expected and large sample sizes needed in its investigation. An equally pessimistic conclusion applies to the JS/Coworkers scale though, as noted in the next section, this results in part from the relatively large proportion of its variance attributable to errors of measurement.

The gain score variances of the QoE/Total, JS/Challenge, and QoE/Challenge scales are markedly reduced from their respective variances as ordinary, between-subjects time two variances, because of their relatively high correlations with themselves over time. Accordingly, an experimental manipulation expected to have the same size of effect if applied in a within-subject, over time design as it would have in a between-subject, single-time design, would be more likely to produce significant results with the within-subject, over-time design. This statistical advantage would be attributable to the smaller error variance of gain scores for the analysis over time. For the other variables studied, however, there is increased variance in the raw gain and gain of standardized scores terms over the between-subjects variance found at time two. Such increased variance is particularly noticeable in the cases of the absences and JS/Coworkers scales as discussed above, and also for the ambiguity variable.

2. Sources of Instability of Measures Over Time

Two major sources of instability in measurements are distinguished. The first is simply errors of measurement, the second "real" change.

(a) Unreliability as a Source of Instability of Measures Over Time. In order to estimate the reliability of a change score, one may start with the separate reliabilities of the measures within each phase (Lord, 1963). The reliability of a measure at any one phase may, for homogeneous scales, be estimated by the internal consistency of the measure. Cronbach's alphas were accordingly computed for the five job satisfaction facets (JS/Resource Adequacy, JS/Financial Rewards, JS/Challenge, JS/Comfort and JS/Coworkers), the global measure of Job Satisfaction, role ambiguity, effort-expended and supervisory rating scales, and appear in Table 4.

The Quality of Employment facets' scales were not designed as homogeneous scales; thus use of Cronbach's alpha for purposes of estimating reliability would be inappropriate. Yet lower bound estimates for the reliability of these scales may be obtained as follows: The QoE measures were originally designed to predict job satisfaction (a measure consisting of both global satisfaction and the five job satisfaction facets) in earlier national survey work (Barnowe, Mangione and Quinn, 1972). As such, the multiple correlation of the QoE facets with job satisfaction was nearly as high as the reliability of the job satisfaction measure permits.

The use of path analytic techniques with unmeasured variables (Duncan,

Table 4

Within Phase Reliabilities and Reliability of Gain Scores

Variable	Reliability at Phase I	Reliability at Phase II	Reliability of Raw Gain[c]	Reliability of gain of Standardized scores[c]
QoE/Resource adequacy	.60[a]	.60[a]	.41	.41
QoE/Financial rewards	.60[a]	.60[a]	.20	.19
QoE/Challenge	.60[a]	.60[a]	.01	.01
QoE/Comfort	.60[a]	.60[a]	.26	.26
QoE/Total	.66[a]	.66[a]	.15	.14
Ambiguity	.54[b]	.55[b]	.35	.35
Global job satisfaction	.73[b]	.71[b]	.48	.48
JS/Coworkers	.49[b]	.46[b]	.30	.30
JS/Challenge	.83[b]	.87[b]	.67	.67
JS/Comfort	.66[b]	.73[b]	.42	.41
JS/Resource adequacy	.66[b]	.67[b]	.38	.38
JS/Financial rewards	.66[b]	.66[b]	.39	.39
Effort expended	.44[b]	.27[b]	-.16	-.17
Supervisor ratings	.92[b]	.94[b]	.88	.88

[a] Lower bound reliability estimates are generated from the data available in Barnowe et al. (1972); See text for details.

[b] Alpha coefficients are computed on present dataset (N=272).

[c] Reliability of gain scores are computed from formulae in Lord (1963, p. 32).

1966; Werts and Linn, 1970b) suggests a way to use the information reported by Barnowe et al. (1972) to obtain a lower bound estimate of the reliability of the QoE facets' linear combination that was generated by the multiple regression of the QoE facets on job satisfaction (which yielded the multiple correlation reported). The two measured variables are job satisfaction and a linear composite of QoE. Measured scores on both variables are functions of their respective "true" scores and errors of measurement. The correlation of the measured QoE composite with measured job satisfaction may be decomposed into three parts (a, b and c). In Figure 1, "a" represents the path coefficient from the "true" QoE composite score to the measured QoE score; "b" is the correlation between the "true" QoE composite score and the "true" job satisfaction score, and "c" the path coefficient from "true" job satisfaction to measured job satisfaction. Thus, a^2 equals the reliability of the composite QoE measure, and c^2 the reliability of the job satisfaction score. As c^2 is known from Barnowe et al. (1972), we are left with two unknowns: "a" and "b". The causal model in Figure 1 assumes uncorrelated errors of measurement (e_1 and e_2) between the measured QoE composite and the measured job satisfaction score.

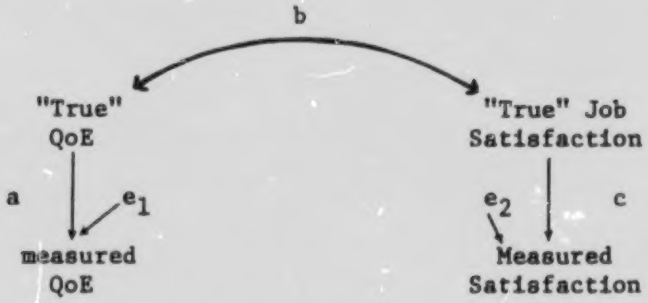

b
"True" QoE
"True" Job Satisfaction
a
e_1
e_2
c
measured QoE
Measured Satisfaction

Figure 1

Given the causal model presented, the correlation between measured QoE and measured job satisfaction equals a x b x c. Clearly "a" will be least when "b" is at its maximum value, as both "c" and the correlation between QoE and job satisfaction are fixed values. Assigning the maximum value to "b" (1.0) yields a lower bound estimate for "a" and, accordingly, for the reliability of the linear QoE composite.

The reliability of this Quality of Employment composite is at least this good (computed to be 0.70), since the model makes the extreme assumption that the correlation between the "true score" on Quality of Employment composite and the "true score" on satisfaction equals one. If that correlation is less than perfect, the reliability of this Quality of Employment composite must exceed the 0.70 computed here for the correlation between the two measures to have been obtained. In order to determine the reliability of the individual facets from the reliability computed for the composite, we can reverse Nunnally's (1967, p.213) formula for obtaining the reliability of a linear composite from the reliability of its components. To do so, however, we must assume that the reliabilities of the separate scales are all equal. The estimate obtained by this procedure of the reliability of each of the Quality of Employment facets (Table 4) is 0.60.

An overall QoE scale also exists that is a composite of the QoE facet scales (QoE/Total). As the facets' weights in this composite differ slightly from those obtained in the Multiple Regression reported above, the correlation of the QoE/Total measure with Barnowe et al.'s (1972) job satisfaction measure was obtained. Assuming again no correlated error, the lower bound estimate (obtained in the same manner as above) for QoE/Total's reliability is 0.66. While Barnowe et al. use the full 33 item measure of QoE/Total, and the present measure uses only 31 of the original 33 items, the two forms intercorrelate very highly (0.99), so the estimate should be insensitive to this difference in item number.

Since many of the items used in the job satisfaction measure have similar phrasing, or cover job facets similar to those of the QoE facets, one might question the assumption here of no correlated error between measures. In the past when this job satisfaction measure has been separated into its facet-specific and facet-free components, greater correlated error might have been expected in the facet-specific case, and hence a higher correlation of facet-specific job satisfaction with the QoE facets. This distinction between facet-specific and facet-free satisfaction has, however, made little difference for the measures' correlations with QoE (Barnowe et al., 1972), so correlated error seems not to be a serious problem.

While lower bounds have been identified for the reliability of QoE facets, more accurate estimates of these reliabilities would probably be obtained if a value more plausible than 1.0 were used for the correlation between true QoE and true job satisfaction. If, for example, 0.80 was used as an arbitrary estimate of the two true scores' intercorrelation, the reliability of the linear composite of QoE becomes 0.87 and the reliability of the individual facets, 0.83. The lower bound estimates will, however, be the ones used, as they offer the more conservative and less arbitrary estimates. Estimates of the reliabilities of the single item measures of involvement, absences, and intent to turnover were not computed.

With the reliability estimates generated within each phase by the procedures used above, the reliabilities of the raw gain and the gain of standardized scores may be computed using the formulae presented in Lord (1963, reproduced in Appendix 2) for those 14 variables for which within-phase reliability estimates were available. These estimates appear in Table 4.

The two analytic approaches based upon an analysis of covariance model (residual and corrected residual scores) in fact use a Phase II measure as the dependent variable, with the Phase I measure as perhaps only one of several covariates. It seems appropriate then to use the Phase II reliability estimates to infer the proportion of variance attributable to errors of measurement for these two approaches to change scores.[1] The proportions of residual and corrected residual variances accounted for by errors of measurement reported are, then, proportions of the total time two variance. In contrast, for raw gain and standardized gain the proportions of variance attributable to errors of measurement are proportions of the variance that remains after the initial differences at time one have been removed (by the computation of the gain scores).

It is clear from Table 4 that the reliabilities of the effort-expended and the JS/Coworkers scales need to be improved, as both have poor within-phase reliabilities and, accordingly, poor change score reliabilities.[2] The negative reliabilities of the change scores of the effort-expended scale result from the fact that the variable correlates with itself over time at a level higher than its within-phases reliabilities would ordinarily permit. These negative reliabilities may be considered to be zero.

---

[1] Formulae for the computation of the reliability of residual gain and corrected residual gain scores appear in Davidson (1972, p.38) and Tucker et al. (1966, pp.468-469) and are reproduced for the reader's convenience in Appendix 2. The interested reader may compute these reliabilities with these formulae from the data appearing in Tables 3 and 4.

[2] A more extensive discussion of the relationships between within-phase reliability, change score reliability and statistical power appears in Fleiss (1976) and in Overall and Woodward (1975, 1976).

The low reliabilities of the JS/Coworkers and effort-expended scales, both as scales and as change scores, may obscure whatever true change exists and are the major sources of variance of these measures. These errors of measurement probably explain the relatively large variances over time found for these variables in the preceding section.

The within-phase reliabilities of the QoE/Challenge, QoE/Comfort and QoE/Financial rewards scales are adequate for research work (lower bound estimates of 0.60) and the lower bound of the reliability of the QoE/Total is slightly higher (0.66). Interestingly, the reliabilities of their change scores (also lower bound estimates, as the within-phase reliabilities used to generate the change score reliabilities are lower bounds) are particularly low, ranging from 0.01 to 0.26. The seeming contradictions between these two sets of results can be resolved by noting that these variables correlate among themselves at a level that is nearly as high as their reliabilities would seem to permit (ranging from 0.45 to 0.60). Accordingly, the removal of the impact of the time one scores on the time two scores leaves little variance of the QoE facet measured in the time two scores, which thus become unreliable measures of the relevant QoE facet. Since no experimental changes in the jobs were introduced in the interim between the Phase I and Phase II measurements for this "control group," there was little change to be measured reliably. It seems likely, then, that there is little spontaneous change in these variables other than measurement error, but clarification of this issue must await precise estimates of the within-phase reliabilities of the QoE facets' measures.

Quite different observations from those made for working conditions may be made regarding many of the worker attitudes and behaviors. Supervisory ratings are an extreme example of this, with very high reliabilities both within phases and as change scores. Substantial spontaneous variation may be expected for this variable that is not a result of errors of measurement. Similar patterns to those observed for supervisory ratings, though less extreme, are found for global (facet-free) job satisfaction and for three of the five job satisfaction facets (JS/Challenge, JS/Financial Rewards, and JS/Comfort).

The results for the other variables are intermediate, with some reliable change and moderate within-phase reliabilities. Notably, the reliabilities of raw gain and of standardized gain for all the 14 variables are very similar. This results in part from the fact that the within-phase variances at times one and two are very similar, so that the equalization of variances within phases by the standardization procedure used for the gain of standardized scores approach scarcely changes the reliability of the change scores. In addition, however, it seems that the reliability of raw gain is fairly insensitive to changes in variances between the two phases. The variance of supervisory ratings, for example, is 1.5 times as great in Phase II as it was in Phase I, yet the reliabilities of raw gain and standardized gain of supervisory ratings are nearly identical.

An overall pattern emerges from these results. By and large, the working conditions scales display relatively little reliable change, whereas a great deal of reliable change is found in worker attitudes and behaviors. Evidently in control group conditions there is little spontaneous change in working conditions but (perhaps because of this) substantial changes in worker attitudes and behaviors occur, possibly as workers react differently to this lack of change in conditions.

(b) Sources of "Real" Change. If unreliability is one source of instability of scores over time, "real" change is the rest. What are the sources of this change, i.e., what variables correlate with change? With the identification of such variables, the spontaneous changes that may be expected to correlate with them but are not related to an experimental manipulation may be controlled statistically. In general, these changes may be considered a result of unplanned or unanticipated changes in work roles and their occupants to which several factors are relevant:

1. People mature in their jobs, gradually adapting themselves to the demands of these jobs. Similarly, events in their lives outside the workplace influence their behavior;

2. Jobs change without any titular change and in ways that are so subtle that they cannot be detected by either Census or Dictionary of Occupational Titles codes;

3. Some people vacate their positions, and others move in to fill them (gross person change in a job);

4. A job may change radically for a person if he or she is transferred, promoted or demoted to a different position (gross job change for a person).

The effects of factors three and four (gross person change and gross job change, respectively) may be easily identified by comparing separately the gross person change group (turnover and replacement subsample of jobs) and the gross job change group (the worker subsample who transferred jobs) with the "no-change" group (workers remaining in the same job). These comparisons are easily performed using dummy variables, where a job is scored "yes" or "no" for person change, and a person is scored "yes" or "no" for having experienced gross job change between phases.

Such a clear operationalization is not possible for factors one and two (spontaneous person change and subtle job change, respectively). Whereas factors one and two are clearly unconfounded with the gross change factors if analysis is performed within the "no-change" group, they are confounded between themselves as no complete and independent measures of either a person changing spontaneously or a job changing subtly are available. The on-the-job observations, which might have fulfilled the role of an independent measure of job conditions have little utility in this regard, since only a subset of job characteristics were reliably and distinctly measurable. It could be argued that, even for this subset, change in job characteristics might have occurred in response to the characteristics of the job holder. Some sense of the importance of these subtle intrapersonal and intrajob changes may be gauged, however, if we realize that these changes are simply weaker forms of gross person change and gross job change. For example, a job may change subtly but such effects are magnified if a job is changed greatly by becoming a completely different job. As such, the impact of factors one and two may be inferred to some extent from the impact of the gross change factors, factors three and four, respectively.

Partial measures of job characteristics (occupational group, income, tenure) and personal demographics (age, education, sex and race), incomplete as they are, will be studied for their impact on scores. Information concerning the distributions of these demographic variables, job characteristics, rate of turnover, and extent of internal (or intraorganizational) mobility is typically available to a researcher before a study begins, so the researcher will know whether to expect score variance that is correlated with these demographic variables, and may then have a way to control statistically for their impact.

In this analysis of sources of change, the dependent variable used will not be any one of the four types of change scores listed by Kenny (1975) and used in the preceding analyses. Rather, the simple unmodified version of the measures themselves is studied. There are two reasons for this choice of the form of the dependent variable. One is that the repeated discussion of a particular correlation, for each of four measurement techniques, would lengthen this report unduly. Furthermore, it is not clear if any of these four change scores (or any other change score for that matter) is particularly suitable for correlational work (Cronbach and Furby, 1970; Davidson, 1972; Kessler, 1977).

The present use of the unmodified measures as the dependent variables will not typically yield conclusive evidence regarding possible sources of change and covariates, but rather will give preliminary information regarding which variables might be appropriate covariates in future work.

The variables studied here include the 17 measures of working conditions, worker attitudes and worker behaviors. The three-item measure of job involvement is used as the Phase II measure of involvement. The sources of change of these measures are job characteristics (occupational group, income and tenure) and personal demographics (age, sex, education and race), as well as dummy variables (gross job change and gross person change). The tenure variable refers to tenure with the company at Phase I and reported tenure on the job at Phase II. In organizations with substantial internal mobility this distinction is important. Here, however, the correlation between the two variables was computed within the subsample of same person in same job to assess the impact of internal mobility prior to Phase I. The correlation was found to be 0.71, indicating a fairly close correspondence in these work sites between company and job tenure.

The relationships between the sources of change variables and the 17 measures of working conditions and work-related attitudes and behavior are studied within a particular phase, since some of the personal demographics associated with a job do change over time (when a job occupant is replaced between the two phases of measurement), and some of the job characteristics likewise change over time (for a person who is transferred to a different job between the two phases of measurement).

Whereas the levels of the demographic variables of age, education, income, and tenure are clearly ordered for use in computing linear coefficients, and the dichotomous variables may enter the analysis as dummy variables with two levels, occupational type is a nominal variable with eight levels. A one way analysis of variance provides an appropriate technique for studying the impact of this nominal variable on the other variables. To ensure reasonably homogeneous categories, the craftsmen and foremen classification was excluded since the number of supervisory duties entailed differs sharply for its two subcategories. Since categories with at least 25 cases were desired for the analysis, the categories for sales people and laborers were also excluded.

In Phase I, then, five occupational categories were used: professionals, managers, clericals, operatives, and service workers. Certain discrepancies were noted in the occupational codes assigned to the same jobs in the two phases. The Phase I codes had been assigned by a very experienced occupational coder (over ten years work) at the Survey Research Center. The coders for Phase II were somewhat less experienced. The Phase II codes were accordingly judged to be less reliable than the Phase I codes, hence only Phase I codes were used in the analysis of both phases. The retention of only Phase I codes meant that workers who had transferred to different jobs by Phase II were excluded from the analysis of variance of the Phase II data, since no codes were available for their Phase II jobs. This depletion in sample size reduced the manager category to below 25 cases, so the category was removed from the Phase II analysis. In Phase II, then, the impact of occupational grouping concerns four occupational categories: professionals, clericals, operatives, and service workers. The results from the analyses of variance appear in the form of eta coefficients that have been added to the bottom row of the correlation matrix in Table 5. The eta resembles Pearson r without the restriction of linearity, and is always positive.

The results of the correlation matrix of Table 5 have been summarized in Tables 6 and 7. Table 6 is a Phase I table, showing all entries which have a significant correlation between one of the nine "sources of change" variables and one of the 17 measures of working conditions, worker attitudes and worker behaviors. Table 7 is the corresponding Phase II table. Since many of these significant correlations do not replicate over the two phases, Table 8 presents those entries that have significant correlations between a "source of change" and a "work" variable in both phases, i.e., replicated correlations.

The intercorrelations of demographic and work-related variables that do not replicate may also be of interest. These demographic variables are particularly likely to be the correlates of change in the dependent variable

Table 5

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

	Age	Sex	Edu-cation	Race	Tenure	Income	Turnover [a] & Replcmt.	Trans-fer/no change [b]	QE/Total	QE/Res-ource	QE/Fin-ancial rewards	QE/Chal-lenge
Phase I												
Age	1.0000											
Sex	.0270 (271)	1.0000										
Education	-.3140* (271)	-.1111 (271)	1.0000									
Race	-.0104 (267)	.2163* (267)	.0823 (266)	1.0000								
Tenure	.6474* (265)	.1373* (266)	-.2133* (265)	.0201 (261)	1.0000							
Income	.1302* (271)	.3796* (271)	.1674* (270)	.3236* (266)	.2215* (264)	1.0000						
Turnover & replacemt. [a]	-.1237 (199)	-.0927 (199)	.0657 (198)	.1379 (195)	-.1634* (194)	-.1574* (198)	1.0000					
Transfer/no change [b]	-.2545* (236)	.0417 (236)	.0032 (235)	.2457* (233)	-.2855* (231)	-.0438 (235)	-0. (163)	1.0000				
Phase I												
QoE/Total	.1570* (271)	.1914* (271)	.1439* (270)	.1451* (266)	.1734* (265)	.3906* (270)	-.1060 (198)	.0074 (235)	1.0000			
QoE/Resource adq.	-.0472 (264)	.0275 (264)	.0732 (264)	-.0082 (259)	-.0660 (258)	.0174 (263)	.0565 (192)	.0413 (229)	.5131* (264)	1.0000		
QoE/Fin. rewards	-.0320 (269)	-.0420 (269)	.2501* (269)	.0348 (264)	-.0727 (263)	.1507* (268)	.0597 (196)	.0151 (233)	.5432* (268)	.1458* (262)	1.0000	
QoE/Challenge	.1573* (271)	.2133* (271)	.1232* (270)	.1692* (266)	.2611* (265)	.4147* (270)	-.1740* (198)	-.0606 (235)	.8594* (271)	.2067* (264)	.3636* (268)	1.0000
QoE/Comfort	.1819* (271)	.0372 (271)	-.0008 (270)	.0598 (266)	.0747 (265)	.2221* (270)	-.0634 (198)	.0926 (235)	.4745* (271)	.2803* (264)	.1747* (268)	.0851 (271)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

	Age	Sex	Education	Race	Tenure	Income	Turnover & replcmt.^{a}	Transfer/no change^{b}	QE/Total	QE/Resource	QE/Financial rewards	QE/Challenge
Phase I												
Global job satisfaction	.2215* (271)	-.0229 (271)	-.0625 (271)	.1096 (266)	.0862 (265)	.1745* (270)	-.1584* (198)	-.1148 (235)	.4670* (270)	.2707* (264)	.1357* (268)	.4103* (270)
JS/Resource adequacy	.0973 (271)	-.0658 (271)	.1174 (270)	.0009 (266)	-.0723 (265)	.0497 (270)	.0248 (198)	.0348 (235)	.5138* (270)	.7250* (264)	.2141* (269)	.2479* (270)
JS/Financial rewards	[illegible]* (270)	-.0193 (270)	-.0082 (270)	.2830* (265)	.0618 (264)	.1407* (269)	.0618 (197)	.0531 (234)	.4193* (269)	.3489* (264)	.2339* (268)	.2868* (269)
JS/Challenge	.2317* (270)	.1178 (270)	.0470 (270)	.0463 (265)	.1419* (264)	.2263* (269)	-.2223* (197)	-.1023 (234)	.6684* (269)	.3689* (264)	.3772* (268)	.5681* (269)
JS/Comfort	[illegible]* (270)	.0006 (270)	.0318 (270)	.0192 (265)	.0599 (264)	.0779 (269)	-.0558 (197)	.0015 (234)	.5155* (269)	.4817* (254)	.2633* (268)	.2457* (269)
JS/Coworkers	[illegible] (271)	.0704 (271)	.0234 (270)	.1082 (266)	-.0095 (265)	.1497* (270)	.0049 (196)	.0072 (235)	.4449* (270)	.3978* (264)	.1721* (269)	.3186* (270)
Ambiguity	[illegible] (271)	-.0745 (271)	-.1367* (270)	-.0949 (266)	-.0460 (265)	-.1505* (270)	.1176 (198)	.0251 (235)	.1571* (271)	.4724* (264)	-.0850 (268)	.0206 (271)
Involvement	[illegible]* ([illegible])	[illegible] (263)	.0125 (267)	.0019 (263)	.2360* (262)	.2359* (267)	-.1085 (195)	-.0596 (234)	.4802* (267)	.1269* (262)	.2248* (266)	.5022* (267)
Effort expended	[illegible] (267)	.1417* (267)	.1690* (266)	.2164* (262)	.2038* (261)	.2533* (266)	-.0343 (194)	.0254 (233)	.1987* (266)	-.0786 (261)	.1117 (265)	.2860* (266)
Intent to turnover	[illegible]* ([illegible])	.1357* (272)	-.1567* (271)	.1923* (267)	.1996* (265)	.1944* (271)	-.2773* (199)	-.0566 (236)	.2160* (271)	.0757 (264)	-.0682 (269)	.2536* (271)
Absences	[illegible] (272)	.0700 (272)	.2033* (271)	.0954 (267)	.1247* (266)	.0685 (271)	-.0287 (199)	.0604 (236)	.1205* (271)	-.0011 (264)	-.0171 (269)	.0878 (271)
Supervisory ratings ^{c}	[illegible] (61)	.2945* (61)	.1424 (61)	.0052 (61)	.1278 (61)	.0629 (61)	-.0696 (50)	.1234 (47)	.2964* (60)	.2745* (59)	.1288 (60)	.2135 (60)
Phase II												
Age	.9775* (263)	.0845 (263)	-.2957* (263)	-.0330 (263)	.6007* (262)	.1499* (267)	-.2093* (195)	-.2475* (233)	.1255* (267)	-.0076 (262)	-.0625 (265)	.1276* (267)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase II	Age	Sex	Edu-cation	Race	Tenure	Income	Turnover[a] & replcmt.	Trans-fer/no change[b]	QE/Total	QE/Res-ource	QE/Fin-ancial rewards	QE/Chal-lenge
Sex	.0530 (263)	.9621* (269)	-.1494* (268)	.2107* (263)	.1631* (262)	.3896* (267)	-.1250 (195)	.0636 (233)	.1687* (267)	.0151 (262)	-.0670 (265)	.2356* (267)
Education	-.2233* (263)	-.0572 (269)	.8640* (263)	.0873 (263)	-.1516* (262)	.1700* (267)	.0155 (195)	.0066 (233)	.1833* (267)	.0214 (262)	.2400* (265)	.1772* (267)
Race	-.0811 (263)	.2415* (263)	.0933 (263)	.9476* (257)	-.0041 (255)	.3418* (259)	.0727 (189)	.2560* (226)	.1373* (259)	.0030 (255)	.0458 (257)	.1472* (259)
Tenure	.4202* (263)	.0731 (263)	-.1353* (268)	-.0696 (263)	.6306* (262)	.0504 (267)	-.3410* (195)	-.3831* (233)	.0200 (267)	-.0436 (262)	-.0906 (265)	.0988 (267)
Income	.0733 (269)	.1395* (269)	.1354* (269)	.0469 (264)	.0834 (263)	.4836* (268)	-.0190 (196)	-.0747 (233)	.1153 (268)	-.0475 (263)	.1037 (266)	.1276* (268)
QoE/Total	.1649* (260)	.0725 (260)	.1624* (260)	.0814 (255)	.2335* (254)	.2747* (259)	-.1313 (187)	-.0544 (225)	.6034* (259)	.2800* (254)	.3495* (257)	.5383* (255)
QoE/Resource adequacy	.1136 (257)	-.0534 (257)	.0217 (257)	.0012 (253)	.0450 (251)	.0427 (256)	-.0096 (186)	.0041 (221)	.3128* (257)	.3228* (252)	.1299* (254)	.2038* (257)
QoE/Financial rewards	-.0334 (255)	-.1053 (255)	.2722* (255)	-.0564 (250)	.0078 (249)	.1014 (254)	.0257 (186)	.0799 (221)	.3644* (254)	.1633* (249)	.5038* (252)	.2598* (254)
QoE/Challenge	.1227* (261)	.1254* (261)	.1831* (261)	.1150 (256)	.2567* (255)	.3002* (260)	-.1615* (188)	-.0872 (225)	.5596* (260)	.1765* (255)	.3127* (258)	.5944* (260)
QoE/Comfort	.2165* (257)	.0329 (257)	-.0193 (257)	.0267 (252)	.1462* (251)	.1897* (256)	-.0909 (184)	.0038 (223)	.3401* (256)	.2123* (251)	.0954 (254)	.1793* (256)
Global job satisfaction	.2837* (264)	-.0213 (268)	-.0640 (268)	.0554 (263)	.2549* (262)	.1605* (267)	-.0863 (195)	-.0156 (233)	.3682* (267)	.2248* (262)	.0980 (265)	.3015* (267)
JS/Resource adequacy	.1524* (267)	.0236 (269)	.0512 (269)	.0769 (264)	.1270* (263)	.0526 (268)	.0062 (196)	.0299 (233)	.3589* (268)	.3642* (263)	.1866* (266)	.2075* (268)
JS/Financial rewards	.2812* (268)	-.0310 (268)	-.0761 (268)	.1456* (263)	.2590* (262)	.1600* (267)	-.1905* (195)	-.0573 (232)	.3690* (267)	.1967* (262)	.1484* (265)	.2974* (267)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase II	Age	Sex	Education	Race	Tenure	Income	Turnover & replcmt.[a]	Transfer/no change[b]	QE/Total	QE/Rescource	QE/Financial rewards	QE/Challenge
JS/Challenge	.2439* (268)	.0416 (263)	.0289 (268)	.0854 (263)	.2485* (262)	.2030* (267)	-.1150 (195)	-.0569 (232)	.4484* (267)	.2566* (262)	.1938* (265)	.4187* (267)
JS/Comfort	.2593* (264)	.0707 (268)	-.0299 (268)	.1081 (263)	.2276* (262)	.1490* (267)	-.0635 (195)	-.0744 (232)	.3795* (267)	.3030* (262)	.0859 (265)	.2671* (267)
JS/Coworkers	.1339* (269)	.0095 (269)	-.0284 (269)	-.0094 (264)	.0675 (263)	.0753 (268)	-.0546 (196)	-.0808 (233)	.1665* (268)	.2012* (263)	.0080 (266)	.1118 (268)
Ambiguity	-.0233 (261)	.0026 (261)	-.0743 (261)	-.0419 (256)	-.0319 (255)	-.1746* (260)	-.1254 (188)	-.0161 (225)	-.0226 (260)	.1935* (255)	-.0369 (258)	-.0724 (260)
Involvement	.2709* (268)	.2238* (268)	-.0595 (268)	.0039 (263)	.2830* (262)	.2780* (267)	-.0973 (195)	-.0615 (233)	.3374* (267)	.1140 (262)	.1900* (265)	.3418* (267)
Effort expended	.0007 (265)	.1120 (265)	.2030* (265)	.0739 (260)	.0668 (259)	.2310* (264)	-.0039 (192)	.0186 (230)	.3114* (264)	.0316 (259)	.1923* (262)	.3419* (264)
Intent to turnover	.2969* (267)	.0236 (267)	-.2182* (267)	.1620* (262)	.2289* (261)	.1191 (266)	-.0460 (194)	-.0495 (232)	.1233* (266)	.0979 (261)	-.1384* (264)	.1120 (266)
Absences	-.0153 (268)	.0101 (268)	.1715* (268)	.0524 (263)	.0577 (262)	.1109 (267)	-.0190 (195)	.0333 (233)	.1693* (267)	.0857 (262)	.1829* (265)	.1154 (267)
Supervisory ratings[c]	.2058 (60)	.3336* (60)	.1169 (60)	.2946* (60)	.2501* (60)	.2505* (60)	-.1033 (50)	.2488 (46)	.2975* (59)	.2518 (59)	.0460 (59)	.3012* (59)
Occupational group[d]	.2572* (236)	.3495* (236)	.5919* (235)	.5838* (232)	.3116* (233)	.5739* (235)	.1728 (175)	.2910* (203)	.4644* (235)	.1946 (231)	.4552* (231)	.5119* (235)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase I

	QE/Comfort	Global job satisfaction	JS/Resource adequacy	JS/Financial rewards	JS/Challenge	JS/Comfort	JS/Coworkers	Ambiguity	Involvement	Effort expended	Intent to turnover	Absences
QE/Comfort	1.0000											
Global job satisfaction	.2555* (270)	1.0000										
JS/Resource adequacy	.3540* (270)	.4233* (270)	1.0000									
JS/Financial rewards	.2814* (269)	.4130* (270)	.4754* (270)	1.0000								
JS/Challenge	.3320* (269)	.5617* (270)	.5220* (270)	.4715* (270)	1.0000							
JS/Comfort	.5243* (269)	.4404* (270)	.6365* (270)	.5477* (270)	.6520* (270)	1.0000						
JS/Coworkers	.2615* (270)	.4031* (270)	.4066* (271)	.3404* (270)	.5198* (270)	.5276* (270)	1.0000					
Ambiguity	.0991 (271)	.1770* (270)	.4039* (270)	.2115* (269)	.2110* (269)	.3244* (269)	.2228* (270)	1.0000				
Involvement	.1576* (267)	.2919* (267)	.1744* (268)	.2341* (267)	.4797* (267)	.2605* (267)	.1473* (268)	.0279 (267)	1.0000			
Effort expended	-.0222 (265)	.0404 (265)	-.0377 (267)	.0509 (266)	.1166 (266)	.0297 (266)	-.0116 (267)	-.2042* (266)	.3412* (265)	1.0000		
Intent to turnover	.1055 (271)	.3894* (271)	.1602* (271)	.2731* (270)	.2490* (270)	.1773* (270)	.2557* (271)	.0403 (271)	.1459* (268)	.0200 (267)	1.0000	
Absences	.1739* (271)	.0786 (271)	.0387 (271)	.0056 (270)	.1531* (270)	.1262* (270)	.0633 (271)	-.0548 (271)	.0907 (268)	.0212 (267)	.0136 (272)	1.0000
Supervisory ratings[a]	.1525 (60)	.1483 (61)	.3157* (61)	.1730 (61)	.3773* (61)	.2763* (61)	.1548 (61)	-.0797 (60)	.2527* (60)	.0636 (60)	.0200 (61)	.3884* (61)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase II	QE/Comfort	Global job satisfaction	JS/Resource adequacy	JS/Financial rewards	JS/Challenge	JS/Comfort	JS/Co-workers	Ambiguity	Involvement	Effort expended	Intent to turnover	Absences
Age	.1657* (267)	.2378* (268)	.0786 (267)	.1853* (267)	.2131* (267)	.1899* (267)	.0748 (267)	.0271 (267)	.2129* (264)	-.0084 (263)	.2433* (268)	.0717 (268)
Sex	.0598 (267)	.0204 (269)	-.0628 (267)	-.0153 (267)	.1399* (267)	.0127 (267)	.1095 (267)	-.0489 (267)	.1043 (264)	.1280* (263)	.1852* (268)	.0746 (268)
Education	.0012 (267)	-.0309 (268)	.0617 (267)	.0016 (267)	.0746 (267)	.0964 (267)	.0258 (267)	-.1570* (267)	.0209 (264)	.2188* (263)	-.1211* (268)	.2136* (268)
Race	.0710 (259)	.0734 (263)	-.0172 (259)	.2189* (259)	.0254 (259)	-.0094 (259)	.1193 (259)	-.1376* (259)	-.0133 (257)	.2042* (256)	.1569* (260)	.0795 (260)
Tenure	-.0339 (267)	.1625* (263)	-.0753 (267)	.0622 (267)	.1291* (267)	.0658 (267)	-.0525 (267)	-.0163 (267)	.1017 (264)	.1129 (263)	.1498* (268)	.0182 (268)
Income	.0714 (263)	-.0546 (269)	-.0177 (268)	-.0909 (268)	.0513 (268)	-.0974 (268)	-.0847 (268)	-.1855* (268)	.0736 (265)	.0983 (264)	-.0667 (265)	.0539 (269)
QoE/Total	.2420* (259)	.3570* (260)	.3605* (259)	.2552* (259)	.5172* (259)	.3534 (259)	.3067 (259)	.0266 (259)	.3608 (256)	.1208 (255)	.1113 (260)	.0908 (260)
QoE/Resource adequacy	.1817* (257)	.3102* (257)	.3990* (256)	.2116* (256)	.3350* (256)	.3333* (256)	.2626* (256)	.1451* (257)	.1263* (253)	-.0708 (252)	.0326 (257)	.0795 (257)
QoE/Financial rewards	.1033 (254)	.1232* (255)	.1854* (254)	.0937 (254)	.2201* (254)	.2401* (254)	.0984 (254)	-.0229 (254)	.1383* (252)	.0690 (250)	-.0695 (255)	.0760 (255)
QoE/Challenge	.0821 (260)	.3020* (261)	.2344* (260)	.1890* (260)	.4415* (260)	.1971* (260)	.2302* (260)	-.0537 (260)	.3807* (257)	.2319* (256)	.0895 (261)	.0623 (261)
QoE/Comfort	.4534* (255)	.2645* (257)	.2895* (256)	.2163* (256)	.3547* (256)	.3455* (256)	.2358* (256)	.1081 (256)	.1761* (253)	-.0933 (252)	.1789 (257)	.0714 (257)
Global job satisfaction	.2407* (267)	.4649* (268)	.2978* (267)	.2490* (267)	.3995* (267)	.3107* (267)	.2620* (267)	.0990 (267)	.3312* (264)	.0928 (263)	.1970* (268)	.1328* (268)
JS/Resource adequacy	.2552* (268)	.3524* (269)	.4623* (268)	.2756* (268)	.4317* (268)	.4574* (268)	.3549* (268)	.1602* (268)	.2247* (265)	-.0003 (264)	.1013 (269)	.1027 (269)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

	QE/Comfort	Global job satisfaction	JS/Resource adequacy	JS/Financial rewards	JS/Challenge	JS/Comfort	JS/Coworkers	Ambiguity	Involvement	Effort expended	Intent to turnover	Absences
Phase II												
JS/Financial rewards	.2423* (267)	.3202* (263)	.3004* (267)	.4417* (257)	.4002* (267)	.3512* (267)	.2579* (267)	.0904 (267)	.2170* (264)	-.0039 (263)	.2163* (268)	.0395 (268)
JS/Challenge	.1655* (267)	.4049* (268)	.3151* (267)	.2751* (267)	.5439* (267)	.3751* (267)	.3742* (267)	.1237* (267)	.4127* (264)	.1382* (263)	.1459* (268)	.0859 (268)
JS/Comfort	.3006* (267)	.3454* (268)	.3640* (267)	.2728* (267)	.4312* (267)	.4816* (267)	.4015* (267)	.1260* (267)	.2613* (264)	.0442 (263)	.1929* (268)	.0937 (268)
JS/Coworkers	.1030 (263)	.2298* (269)	.2201* (268)	.1602* (268)	.2178* (268)	.1735* (268)	.2494* (268)	.1557* (269)	.0646 (265)	-.0288 (264)	.1448* (269)	-.0061 (269)
Ambiguity	.0721 (260)	.0718 (261)	.1535* (260)	.0369 (260)	.0976 (260)	.1475* (260)	.1008 (260)	.3048* (260)	-.0594 (257)	-.1370 (256)	-.0664 (261)	.0351 (261)
Involvement	.0935 (267)	.2420* (268)	.2088* (267)	.2379* (267)	.[illegible]089* (267)	.1311* (267)	.0947 (267)	.0218 (267)	.4075* (264)	.2392* (263)	.1645* (268)	.0814 (268)
Effort expended	.0379 (264)	.1118 (265)	.1258* (264)	.1629* (264)	.2448* (264)	.1560* (264)	.1175 (264)	.0314 (264)	.2720* (261)	.4469* (260)	-.0415 (265)	.0681 (265)
Intent to turnover	.1807* (265)	.2919* (267)	.1328* (265)	.2380* (266)	.1498* (266)	.2033* (266)	.1879* (266)	.1171 (266)	.1160 (263)	-.0608 (262)	.3710* (267)	.0470 (267)
Absences	.0923 (267)	-.0313 (268)	.1060 (267)	.0211 (267)	.0757 (267)	.0337 (267)	.0729 (267)	-.0841 (267)	.1336* (264)	.1000 (263)	.0106 (268)	.1348* (268)
Supervisory ratings [c]	.0353 (59)	.2313 (60)	.3306* (60)	.3965* (60)	.4126* (60)	.3143* (60)	.0744 (60)	.0354 (59)	.2843* (59)	.1900 (59)	.3479* (60)	.0851 (60)
Occupational group [d]	.1238 (235)	.1385 (235)	.1804 (235)	.1527 (234)	.3566* (234)	.2062* (234)	.2023* (235)	.3082* (235)	.3563* (233)	.2958* (232)	.2279* (236)	.1904 (236)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase I												
Supervisory ratings[c]	1.0000											
Phase II												
Age	.2721* (60)	1.0000										
Sex	.2854* (60)	.0974 (268)	1.0000									
Education	.0601 (60)	-.2799* (268)	-.0866 (268)	1.0000								
Race	-.0172 (59)	-.0064 (260)	.2369* (260)	.0740 (260)	1.0000							
Tenure	.2487 (60)	.4769* (268)	.0763 (268)	-.1219* (268)	-.0557 (260)	1.0000						
Income	.1812 (61)	.0675 (268)	.1414* (268)	.1643* (268)	.0673 (260)	.0142 (268)	1.0000					
QoE/Total	.4441* (58)	.1553* (260)	.0531 (260)	.1619* (260)	.0817 (252)	.0806 (260)	.1138 (260)	1.0000				
QoE/Resource adequacy	.2603* (58)	.0471 (256)	-.0650 (256)	.0858 (256)	-.0060 (249)	.0152 (256)	-.0432 (257)	.6332* (256)	1.0000			
QoE/Financial rewards	-.0403 (56)	-.0300 (255)	-.1246* (255)	.2624* (255)	-.0641 (248)	-.0752 (255)	.1156 (255)	.4975* (247)	.1850* (243)	1.0000		
QoE/Challenge	.3754* (59)	.1272* (260)	.0987 (260)	.1601* (250)	.1253* (252)	.1065 (260)	.1595* (261)	.9052* (260)	.4023* (257)	.3513* (247)	1.0000	
QoE/Comfort	.2927* (56)	.2234* (257)	.0470 (257)	-.0163 (257)	.0069 (249)	.0245 (257)	.0504 (257)	.5778* (256)	.3012* (252)	.1583* (244)	.3144* (256)	1.0000
	Super. ratings[c]	PII Age	PII Sex	PII Education	PII Race	PII Tenure	PII Income	PII QE/Total	PII QE/Resource adequacy	PII QE/Financial rewards	PII QE/Challenge	PII QE/Comfort

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

	Super. ratings[a]	PII Age	PII Sex	PII Education	PII Race	PII Tenure	PII Income	PII QE/Total	PII QE/Resource adequacy	PII QE/Financial rewards	PII QE/Challenge	PII QE/Comfort
Phase II												
Global job satisfaction	.2400 (60)	.2815* (268)	.0104 (268)	-.0560 (268)	.0561 (260)	.1613* (268)	.0466 (268)	.6046* (260)	.4059* (256)	.2178* (255)	.5212* (260)	.4058* (257)
JS/Resource adequacy	.2857* (61)	.0865 (268)	.0098 (268)	.0794 (268)	.0836 (260)	.0149 (268)	-.0183 (269)	.6377* (260)	.7737* (257)	.2783* (255)	.4505* (261)	.3542* (257)
JS/Financial rewards	.2014 (61)	.3108* (267)	-.0199 (267)	-.0957 (267)	.1720* (259)	.1870* (267)	.0779 (268)	.6050* (259)	.3732* (256)	.2998* (254)	.5006* (260)	.4553* (256)
JS/Challenge	.3351* (61)	.2169* (267)	.0377 (267)	.0342 (267)	.0974 (259)	.1252* (267)	.0953 (268)	.7498* (259)	.5257* (256)	.2950* (254)	.6944* (260)	.3792* (256)
JS/Comfort	.1658 (61)	.2525* (267)	.0573 (267)	-.0111 (267)	.1123 (259)	.1081 (267)	.0590 (268)	.6112* (259)	.5066* (256)	.2296* (254)	.4424* (260)	.5470* (256)
JS/Coworkers	.1155 (61)	.0872 (268)	.0077 (268)	-.0316 (268)	-.0071 (260)	.0838 (268)	.0018 (269)	.4066* (260)	.4692* (257)	.1053 (255)	.3343* (261)	.1840* (257)
Ambiguity	-.0653 (57)	.0279 (260)	-.0226 (260)	-.0664 (260)	-.0092 (252)	.0370 (260)	-.1197 (261)	.2740* (260)	.4870* (257)	.0679 (247)	.1524* (261)	.1171 (256)
Involvement	.2330 (60)	.2568* (268)	.2415* (268)	-.0611 (268)	-.0104 (260)	.1475* (268)	.1342* (268)	.3333* (260)	.0577 (256)	.1026 (255)	.4212* (260)	.1057 (257)
Effort expended	.3637* (58)	-.0316 (265)	.0941 (265)	.1930* (265)	.0312 (257)	.0520 (265)	.1595* (265)	.2193* (257)	.0053 (253)	.1508* (252)	.3139* (257)	.0035 (254)
Intent to turnover	.0010 (59)	.3302* (267)	.0565 (267)	-.2325* (267)	.1562* (259)	.1662* (267)	.0410 (267)	.1755* (259)	.1192 (255)	-.0958 (254)	.1395* (259)	.2420* (256)
Absences	.2507* (60)	-.1104 (268)	.0130 (268)	.1426* (268)	.0449 (260)	-.0201 (268)	.0921 (268)	.2159* (260)	.1446* (256)	.2575* (255)	.1375* (260)	.1851* (257)
Supervisory ratings[a]	[illegible]* ([illegible])	.1892 (59)	.3945* (59)	.1740 (59)	.2474 (58)	.2220 (59)	.3296* (60)	.4380* (57)	.4080* (57)	-.0476 (55)	.3431* (58)	.1770 (55)
Occupational group[d]	.2702 (60)	.2545* (149)	.3299* (149)	.6724* (149)	.5436* (146)	.2561* (149)	.1256 (150)	.2905* (142)	.1640 (142)	.3120* (142)	.3870* (142)	.0962 (139)

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

	PII Global job satisfaction	PII JS/ Resource ade.	PII JS/ Financial rewards	PII JS/ Challenge	PII JS/ Comfort	PII JS/ Coworkers	PII Ambig- ity	PII Involve- ment	PII Effort expended	PII Intent to turnover	PII Absences	PII Super. ratings^c
Phase II												
Global job satisfaction	1.0000											
JS/Resource adequacy	.5046* (268)	1.0000										
JS/Financial rewards	.5018* (267)	.4588* (268)	1.0000									
JS/Challenge	.6735* (267)	.6773* (268)	.5633* (268)	1.0000								
JS/Comfort	.5644* (267)	.6605* (268)	.5926* (268)	.6827* (268)	1.0000							
JS/Coworkers	.3505* (268)	.4451* (269)	.2839* (268)	.5129* (268)	.4589* (268)	1.0000						
Ambiguity	.1513* (261)	.4035* (261)	.2361* (260)	.2676* (260)	.3703* (260)	.3329* (261)	1.0000					
Involvement	.3344* (268)	.1785* (268)	.2655* (267)	.3719* (267)	.2134* (267)	.1552* (265)	-.1045 (260)	1.0000				
Effort expended	.0598 (265)	.0541 (265)	.0283 (264)	.1876* (264)	.0467 (264)	.0203 (265)	-.0458 (257)	.3392* (265)	1.0000			
Intent to turnover	.3238* (267)	.1758* (267)	.2865* (266)	.2759* (266)	.2848* (266)	.1556* (267)	.0623 (259)	.0534 (267)	-.0658 (264)	1.0000		
Absences	.1052 (268)	.1308* (268)	.1312* (267)	.0984 (267)	.0791 (267)	.0203 (268)	.0414 (260)	.0501 (268)	.1585* (265)	-.0367 (267)	1.0000	
Supervisory ratings^c	.3816* (59)	.4694* (60)	.4291* (60)	.4571* (60)	.2983* (60)	.1376 (60)	.1277 (58)	.2847* (59)	.3027* (57)	.0556 (58)	.0424 (59)	1.0000

Table 5 (continued)

Correlation Matrix of Phase I and Phase II Variables (N's appear in parentheses)

Phase II	PII Global satisfaction	PII JS/ Resource adequacy	PII JS/ Financial rewards	PII JS/ Challenge	PII JS/ Comfort	PII JS/ Coworkers	PII Ambiguity	PII Involvement	PII Effort expended	PII Intent to turnover	PII Absences	PII Super. rating[c]
Occupational group[d]	.0970 (149)	.1510 (150)	.0781 (149)	.1457 (149)	.0289 (149)	.0900 (150)	.0853 (143)	.0908 (149)	.2725* (146)	.2574* (148)	.1314 (149)	.1418 (44)

---

* $p < 0.05$

[a] This variable compares turnover and no change jobs, excluding people who transferred jobs between the two phases.

[b] This variable compares transferred jobs to people remaining in the same job, excluding jobs that experienced turnover between the two phases.

[c] Supervisory ratings available in both phases for only 61 cases.

[d] Eta coefficients are reported from analyses of variance. Sample limited to occupational categories with 25 or more cases.

Table 6

Significant Associations of Phase I Demographics with Phase I Work-related Variables

Work-related variables	Age	Sex	Education	Race	Income	Tenure	Occupational group	Turnover[a]	Transfer[b]
Ambiguity			*		*		*		
Involvement	*				*	*	*		
Effort expended		*	*	*	*	*	*		
Intent to turnover	*	*	*	*	*	*	*	*	
JS/Coworkers							*		
JS/Challenge	*	*			*	*	*	*	
JS/Comfort	*						*		
JS/Resource adequacy									
JS/Financial rewards	*			*	*		*		
QoE/Total	*	*	*	*	*	*	*	*	
QoE/Resource adequacy									
QoE/Financial rewards			*		*		*		
QoE/Challenge	*	*	*	*	*	*	*		
QoE/Comfort	*				*			*	
Global job satisfaction	*				*			*	
Supervisor rating		*							
Absences			*			*			

* Significant association found, $p < 0.05$

a This variable compares workers who subsequently turned over with workers remaining in their same jobs, excluding workers who were subsequently transferred.

b This variable compares workers who subsequently transferred with those remaining in the same jobs, excluding workers who subsequently turned over.

Table 7

Significant Associations of Phase II Demographics with Phase II Work-related Variables

Work-related variables	Age	Sex	Education	Race	Income	Tenure	Occupational group	Replacement[a]	Transferred[b]
Ambiguity					*				
Involvement	*	*			*	*			
Effort expended		*	*		*		*		
Intent to turnover	*		*	*		*	*		
JS/Coworkers									
JS/Challenge	*					*			
JS/Comfort	*								
JS/Resource adequacy									
JS/Financial rewards	*			*		*		*	
QoE/Total	*			*			*		
QoE/Resource adequacy									
QoE/Financial rewards			*				*		
QoE/Challenge	*		*	*	*		*	*	
QoE/Comfort	*								
Global job satisfaction	*					*	*		
Supervisor rating		*			*				
Absences			*						

* Significant associations found, $p < 0.05$

[a] Replacement workers are compared with the workers who remained in their jobs, excluding workers who had transferred jobs between Phases I and II.

[b] Transferred workers are compared with workers who remained in their jobs, excluding jobs whose occupants had been replaced by Phase II.

Table 8

Significant Associations of Demographic and Work-related Variables that were Replicated in Both Phases

Work-related variables	Age	Sex	Education	Race	Income	Tenure	Occupational group
Ambiguity							
Involvement	*				*	*	
Effort expended		*	*		*		*
Intent to turnover	*		*	*		*	*
JS/Coworkers							
JS/Challenge	*					*	
JS/Comfort	*						
JS/Resource adequacy							
JS/Financial rewards	*			*			
QoE/Total	*			*			*
QoE/Resource adequacy							
QoE/Financial rewards			*				*
QoE/Challenge	*		*	*	*		*
QoE/Comfort	*						
Global job satisfaction							
Supervisor rating		*					
Absences			*				

* Association significant in both phases at .05 level.

according to arguments outlined by Davidson (1972). Unfortunately, owing to the different correlations in the two phases, these demographics are not likely to be useful as covariates in designs using a raw gain or standardized gain change score, since the differing correlations will likely yield different slopes of the covariate on the dependent variable, in the time one and time two groups. This lack of homogeneity of slopes violates an important assumption governing the interpretation of the results of the analysis of covariance. Nonetheless, such demographics may be used in designs where the residual gain and corrected residual gain scores are appropriate, since only the time two score is used as a dependent variable.

Within the limits imposed by the caveats cited above, the results presented in Tables 5 to 8 can be useful to future researchers. For purposes of the present study, the impact of the variables of gross job change and of gross person change is of particular concern, as they correspond to the factors three and four, postulated earlier. The effects of gross job change (transfer) can be seen in the summary table for Phase II to have had no significant impact on the mean of any of the 17 measures. Since the means of the two groups of people were also not significantly different prior to the transfer of the transfer cases, the Phase II measures are entirely appropriate for measuring the impact of job transfer and change scores are therefore unnecessary (Cronbach & Furby, 1970). As an experimental manipulation, then, non-systematic gross job change does not seem to be a direct source of score variance.

Gross person change in a job, by turnover and replacement, is significantly correlated with two of the 17 Phase II work-related measures; the replacement workers are less satisfied than their continuing co-workers with their financial rewards and they experience less challenge on their jobs.

Initial differences in Phase I for these jobs existed among these two variables for only the QoE/Challenge measure. Other differences were found only at Phase I between the workers who were to leave their jobs and the continuing workers. Certain of the work-related measures, then, are correlated with turnover and replacement in the organization.

Aside from the effects of particular demographic variables on the work-related measures, some general trends may be noted. Among the demographic variables studied here, age seems to have consistent effects on the widest range of variables: of the 17 variables, age replicates its significant correlation with eight over the two phases. The other demographic variables replicate significantly for five or fewer variables. While any of these specific relationships may hold interest for particular researchers, it would seem that age should usually be controlled.

Prime candidates as sources of change would seem to be income, tenure, and occupational group, as these have significant correlations in one phase that do not replicate in the other phase with (respectively) eight, six, and six of the work-related variables. As explained above, the usefulness of these sources of change variables as covariates is likely to be limited to those designs suitable for the use of a residual or corrected residual gain score.

Of the 17 job variables studied, two seem particularly "predictable" by demographic variables: QoE/Challenge and reported intention to look for a new job. Researchers studying these variables may wish to control for the impact of demographic variables upon them, and may choose to include demographic variables explicitly in the theoretical analyses of these variables.

The intercorrelations between the other variables measured in the study (worker perceptions, attitudes and behaviors) should also help future researchers to identify suitable covariates (cf Table 5). These covariates would typically be measured at Phase I (i.e., before the experimental manipulation) to avoid contamination of the covariates by the experimental manipulation. Thus the relevant correlations for the future researcher are those between the Phase I and the Phase II measures.

The usefulness of these variables as covariates, however, would likely be restricted to those designs using residual or corrected residual gain scores, since if there is a significant correlation over an extended lag, this correlation is probably still larger with zero lag, i.e., within Phase I. Differing correlations of the variable with the dependent variable for the two phases would likely produce differences between the phases in the covariate-dependent variable slopes. This would rule out the analysis of covariance for designs suited to a raw gain or standardized gain change score.

In summary, then, the four sources of variance (gross job change, gross person change, spontaneous person change and subtle job change) have different impacts on the measures of working conditions and work-related attitudes and behaviors. Gross job change seems to have no direct impact on any measure. Gross person change (i.e., turnover and replacement) is correlated with certain of the variables studied. Of the personal characteristics studied, age seems to have wide ranging effects on work-related measures. Of the job-related demographics studied, some correlated variance with the work-related variables may be observed, which may extend to the variance of changes in the work measures as well, particularly for tenure, income and occupational group.

As dependent variables, QoE/Challenge and intent to turnover seem particularly "predictable" by demographics and job characteristics, and their variance is likely to be reduced by the use of covariates in their analysis.

3. Identifying Situations of Varying Stability of Scores

The preceding analysis focused on the direct impact of gross job change, gross person change and various demographic variables on work-related measures. Identification of such impact enables a researcher to use statistical control techniques on a given sample.

There is another kind of source of variance not amenable to linear statistical control techniques that is most appropriately handled by the selection of a suitable research population. This kind of variance is the variance associated with being at a particular level of a demographic variable, i.e., the amount of variance within a level of some variable's distribution, rather than the variance correlated with the difference between levels of a demographic distribution. The issue of heteroscedasticity being discussed here concerns different variances within subpopulations, i.e., within the levels of another variable.

It is not difficult to assess the amount of variance over time associated with gross job change and with gross person change as compared to no change conditions. Increased variance associated with such gross changes indicates an upper limit to the increased instability over time expected to result from the subtler person and job changes referred to in the preceding section as factors one and two. The information resulting from these analyses may show which variables are likely, for a particular form of change score, to have increased variances over time, and thus enable the researcher to know whether organizational conditions of high turnover (gross person change) or high internal mobility (gross job change) are likely to increase the error variance term in a projected study.

Since information concerning the demographic distribution of a sample is typically available, the identification of naturally "unstable" populations is also likely to be valuable. A less "unstable" population is, of course, not necessarily the optimal one for research purposes, despite the seeming advantages of a small error term. The very lack of change over time may mean that an "invariant" population is a relatively inflexible population, not easily modified by either natural events or experimental manipulation. Fuller interpretation of such heteroscedasticity is, however, beyond the scope of this paper.

The present analysis of the variance of change scores encompasses all four measurement techniques (raw gain, gain of standardized scores, residual and corrected residual scores). The possible sources of heteroscedasticity include gross job change, gross person change, income, occupational group, race, sex, tenure, age, and education.

While these variables are typically confounded (e.g., in Table 5 for this sample), cross-classification of subjects into distinct subcategories, which would enable analysis of the separate effects, as well as interaction-type effects, would reduce the sample sizes to such an extent that comparisons of the variances in the subcategories would be pointless.

Gross Job Change. Estimates of the four variance scores over time were computed for the job transfer and no-change groups, using the 17 measures of working conditions, worker attitudes and behaviors. Variance of corrected residual scores could not be computed for the three one item measures of involvement, absences and intent to turnover, since no reliability estimates were available for them. There is no obvious reason to expect increased variance of change scores when job conditions remain constant. Consequently, unidirectional statistical hypotheses were tested, based on the prediction that changes in jobs would, if anything, increase the variance of change scores relative to the no-change condition. The increase in variance due to gross job change was tested by constructing F ratios for each variable, with each type of change score, by dividing the variance over time for the transfer sample by the appropriate variance over time within the no-change group. The standard .05 level of significance was chosen.

Significant increases in variance over time were found for one or more of the change score types for five variables: QoE/Resource adequacy, QoE/Challenge, QoE/Total, involvement and JS/Coworkers. The different variance over time scores all depend on one or more of the following components: variances at time one, variances at time two, the cross-phase correlations of the variable with itself, and the Phase I reliability. Since different estimates are sensitive to different components, a significant increase for one change score type does not necessarily imply that they all change. All estimates prove sensitive to the cross-phase correlation of the variable with itself, and all but variance of standardized gain are sensitive to variance at time two.

The greater variance within time two of the scores of the transfer sample may explain the significant increases in variance over time for QoE/Resource adequacy, QoE/Challenge, QoE/Total and involvement, and the absence of significant increases for the standardized gain score for the latter three of these four variables.

The variances of all four gain scores increased significantly for QoE/Resource adequacy.

For QoE/Challenge, as might be expected from the discussion above, the only gain score variance that does not increase significantly under job change is standardized gain ($F_{72,150} = 1.32$, $p < 0.078$).

For QoE/Total, the variance of standardized gain scarcely increases at ($F_{72,150} = 1.05$, $p < 0.389$). The variances of residual and corrected residual scores increase significantly, and the increase of variance of raw gains attains borderline significance ($F_{72,150} = 1.36$, $p < 0.061$).

For involvement, the increase in variance of standardized gain is not significant ($F_{72,157} = 1.14$, $p < 0.250$). The increase for residual gain is significant, but not for raw gain ($F_{72,157} = 1.28$, $p < 0.104$).

The increased variance among transfer workers at time two is not found for JS/Coworkers. The scale for JS/Coworkers has, however, a relatively low Phase I reliability (0.49) which may explain why the only change score of JS/Coworkers that increases significantly under gross job change is the corrected residual score ($F_{72,158} = 1.41$, $p < 0.039$).

Gross Person Change. In the same manner as described above, the variance estimates over time were computed within levels of the gross person change variable: the turnover and replacement group, and the no change group. F ratios were constructed to test the significance of the increase in variances attributable to gross person change, and the standard (one-tailed) .05 level of significance of a standard F table was again used. The unidirectional hypothesis that changing the person will, if anything, increase the variance of changes relative to constant conditions and people, is tested.

Significant increases were found for one or more change score types for all variables except QoE/Comfort and involvement. The variances over time and their components appear in Table 9. These significant increases are due in part to the increased variance (particularly the measures of working conditions) at time two for the replacement workers, but are also due to the striking reductions in the cross-phase correlations of 13 of the 17 variables among the turnover and replacement sample. An increased error variance term for organizations with high turnover would seem to be the rule here, though for certain variables and change scores this increase does not apply, as can be seen in Table 9.

Job Characteristics. The effects over time of the job characteristic variables of occupational group and income are meaningfully studied only within jobs, that is, jobs that presumably retain these characteristics. Accordingly, the group of people who had transferred jobs between phases of measurement are excluded from this analysis.

There were two levels to the income variable: above and below the median reported income at time one. There were four levels of the occupational group variable, retaining only those categories with 25 or more cases: professionals, clericals, operatives, and service workers.

Two-tailed significance tests were used for the F ratios constructed to test the significance of differences in variance between the above and below median income groups. Thus, F's corresponding to a 0.025 level in a typical F table were used to achieve a (two-tailed) 5% significance level. Bartlett's test for homogeneity of variance was used to test the homogeneity of variances for the four category occupational group variable.

The F ratios revealed significant differences in variances over time

Table 9

Tests for Homogeneity of Variances Over Time for Turnover and No Change Samples

	Variance at time one	Variance at time two	$r_{1,2}$	Reliability at time one	df^b	F, variance of raw gain	p^f	F, variance of standardized gain	p^f	F, variance of residual gain	p^f	F, variance of corrected residual gain	p^f
QoE/Resource adequacy													
Turnover	0.31	0.82	.09	.60	34		0.011		0.014		0.003		0.012
No change	0.58	0.53	.47		146	1.78		1.73		1.98		1.76	
QoE/Financial rewards													
Turnover	0.40	0.77	-.06	.60	33	3.00	0.000	2.71	0.000	2.16	0.001	1.71	0.016
No change	0.48	0.56	.61		148								
QoE/Challenge													
Turnover	0.59	0.77	.09	.60	35	4.80	0.000	3.27	0.000	3.43	0.000	2.32	0.000
No change	0.59	0.56	.72		150								
QoE/Comfort													
Turnover	0.29	0.39	.35	.60	33	1.30	0.146	1.21	0.217	1.31	0.140	1.24	0.193
No change	0.31	0.33	.46		148								
QoE/Total													
Turnover	0.20	0.47	.31	.66	34	2.98	0.000	2.05	0.002	3.11	0.000	2.65	0.000
No change	0.23	0.24	.66		150								
Global job satisfaction													
Turnover	1.35	1.20	.22	.73	34	2.29	0.000	1.60	0.029	1.67	0.019	1.61	0.028
No change	0.86	0.92	.51		159								
JS/Coworkers													
Turnover	0.28	0.49	.17	.49	35	1.13	0.297	1.02	0.451	1.81	0.008	1.82	0.010
No change	0.40	0.27	.16		158								
JS/Challenge													
Turnover	0.54	0.48	.21	.83	35	2.53	0.205	0.01	0.399	1.79	0.020	1.76	0.018
No change	0.37	0.38	.58		157								
JS/Comfort													
Turnover	0.21	0.44	.48	.66	35	1.22	0.153	1.05	0.130	1.65	0.023	1.67	0.038
No change	0.28	0.26	.47		157								

Table 9 (continued)

Tests for Homogeneity of Variances Over Time for Turnover and No Change Samples

	Variance at time one	Variance at time two	$r_{1,2}$	Reliability[a] at time one	df[b]	F, variance of raw gain	p[f]	F, variance of standardized gain	p[f]	F, variance of residual gain	p[f]	F, variance of corrected residual gain	p[f]
JS/Resource adequacy													
Turnover	0.26	0.58	.36	.66	35	1.28	0.000	1.32	0.000	1.63	0.001	1.55	0.004
No change	0.47	0.42	.51		158								
JS/Financial rewards													
Turnover	0.47	0.67	-.03	.66	35	2.85	0.000	2.24	0.001	2.10	0.018	1.89	0.042
No change	0.44	0.45	.54		157								
Ambiguity													
Turnover	0.28	0.48	-.24	.54	35	2.28	0.013	2.11	0.007	1.68	0.051	1.53	0.465
No change	0.37	0.32	.41		150								
Effort expended													
Turnover	0.47	0.60	.05	.44	35	1.75		1.86		1.51		1.01	
No change	0.61	0.52	.49		153								
Supervisory rating													
Turnover	0.99	1.03	-.11	.92	13[c]	1.75	0.094	2.69	0.010	<1		<1	
No change	0.98	1.92	.59		35[c]								
Absences													
Turnover	7.52	14.02	.14	*d*	34	<1		<1		1.55	0.038	*e*	
No change	18.11	8.96	.11		159								
Intent to turnover													
Turnover	3.19	1.96	.36	*d*	34	1.71	0.015	<1		1.11	0.327	*e*	
No change	1.24	1.76	.35		158								
Involvement													
Turnover	0.93	0.71	.30	*d*	32	1.48	0.062	1.27	0.169	1.22	0.213	*e*	
No change	0.75	0.66	.45		157								

[a] Previously obtained reliability estimates.
[b] Changing degrees of freedom reflect missing data.
[c] The small number of degrees of freedom is because supervisory ratings were obtained for only 61 cases for both phases.
[d] No reliability estimates available.
[e] Corrected residuals cannot be computed because reliability estimates are unavailable.
[f] Probabilities reported are exact, rounded to the nearest thousandth.

for the different income level groups, for 10 of the 17 measures. The significant F ratios for these 10 variables are reported in Table 10. For all but two variables (QoE/Resources adequacy and JS/Resource adequacy), the greater variance over time was found in the below-median income group.

Six of these 10 variables display heteroscedasticity for raw gain scores, and for four of these, raw gain is the only gain score that has income-related variance differences. Five of these six variables are attitudinal or behavioral measures, the other being QoE/Challenge.

Three of the four variables that do not display differences in variance of raw gain scores are measures of working conditions (QoE/Total, QoE/Comfort and QoE/Resource adequacy). Three of these same four variables also display differences only for variance of standardized gain (QoE/Total, QoE/Resource adequacy and JS/Resource adequacy).

It would seem, then, that there are definite differences in the amount of variance over time to be expected between lower and higher income workers, where most commonly there is more variance over time among the lower income workers. Attitudinal and behavioral measures are most likely to have income-related differences in raw-gain score variance, whereas measures of working conditions are relatively more likely to have income-related differences in the variance of standardized gains.

Occupational group was found to be a source of heteroscedasticity for at least one gain score for all work-related variables except global job satisfaction (see Table 11). Among these 16 variables, significant heteroscedasticity was found for 12 variables for raw gains, for 11 for residual gains, for 10 for standardized gains and for eight variables for corrected residual gain score variance. It is notable that in 30 of the 41 signifi-

Table 10

Homogeneity of Variance Tests for Variables with Significant Differences in Variances over Time between the Above and Below Median Income Groups

	df[a]	Variance of Raw Gain	F[b]	p[c]	Variance of Standardized Gain	F[b]	p[c]	Variance of Residual Gain	F[b]	p[c]	Variance of Corrected Residual Gain	F[b]	p[c]
QoE/Resource Adequacy													
Low income	97	0.59	1.30	0.214	0.97	1.58	0.030	0.43	1.29	0.228	0.50	1.14	0.532
High income	82	0.77			1.54			0.55			0.57		
QoE/Challenge													
Low income	97	0.65	1.56	0.036	1.16	1.91	0.002	0.45	1.20	0.390	0.50	1.08	0.714
High income	87	0.42			0.61			0.38			0.54		
QoE/Comfort													
Low income	95	0.42	1.37	0.140	1.11	1.04	0.868	0.36	2.01	0.002	0.40	2.03	0.002
High income	85	0.30			1.15			0.18			0.20		
QoE Total													
Low income	96	0.25	1.42	0.100	1.07	1.68	0.014	0.22	1.36	0.142	0.23	1.19	0.412
High income	87	0.18			0.64			0.16			0.20		
Global Job Satisfaction													
Low income	101	1.37	1.80	0.004	1.22	1.26	0.270	0.90	1.48	0.058	0.93	1.44	0.074
High income	91	0.76			0.97			0.61			0.64		
JS/Coworkers													
Low income	102	0.72	1.68	0.012	1.82	1.25	0.288	0.29	1.11	0.618	0.29	1.19	0.394
High income	90	0.43			1.46			0.32			0.34		

(continued. . . . .)

Table 10 continued

Homogeneity of Variance Tests for Variables with Significant Differences in Variances over Time between the Above and Below Median Income Groups

	df[a]	Variance of Raw Gain	F[b]	p[c]	Variance of Standardized Gain	F[b]	p[c]	Variance of Residual Gain	F[b]	p[c]	Variance of Corrected Residual Gain	F[b]	p[c]
JS/Resource Adequacy													
Low income	102	0.38	1.45	0.070	0.85	1.49	0.050	0.29	1.37	0.126	0.33	1.26	0.260
High income	90	0.55			1.27			0.40			0.42		
JS/Financial Rewards													
Low income	101	0.70	1.52	0.042	1.32	1.25	0.278	0.47	1.27	0.250	0.49	1.22	0.338
High income	90	0.46			1.06			0.37			0.40		
Absences													
Low income	101	28.59	1.63	0.018	1.97	1.35	0.142	13.26	2.38	0.000	—[d]		
High income	91	17.56			1.46			5.58					
Intent to Turnover													
Low income	101	2.81	1.61	0.022	1.36	1.06	0.674	1.75	1.25	0.276	—[d]		
High income	90	1.75			1.28			1.40					

[a] degrees of freedom change due to missing data

[b] F's always are computed by dividing the larger variance by the smaller

[c] Probabilities are exact, two-tailed probabilities rounded to the nearest thousandth

[d] No reliability estimates available, so corrected residual scores can not be computed.

Table 11

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	p^a
QoE/Resource Adequacy						
Variance of:						
Raw gain	0.82	0.54	0.91	0.61	6.55	0.088
Residual gain	0.60	0.32	0.58	0.50	7.19	0.066
Standardized gain	1.89	0.68	1.48	1.35	17.03	0.001
Corrected residual gain	0.60	0.44	0.59	0.53	2.15	0.543
QoE/Financial Rewards						
Variance of:						
Raw gain	1.47	0.31	0.42	0.52	55.72	0.000
Residual gain	0.92	0.20	0.37	0.44	39.55	0.000
Standardized gain	2.08	1.37	0.80	1.05	22.05	0.000
Corrected residual gain	0.92	0.21	0.47	0.49	33.04	0.000
QoE/Challenge						
Variance of:						
Raw gain	0.63	0.35	0.65	0.85	13.18	0.004
Residual gain	0.51	0.20	0.44	0.58	18.55	0.000
Standardized gain	1.29	0.74	1.10	1.25	6.00	0.112
Corrected residual gain	0.54	0.26	0.49	0.62	12.61	0.006
QoE/Comfort						
Variance of:						
Raw gain	0.37	0.28	0.40	0.50	5.42	0.144
Residual gain	0.24	0.28	0.25	0.40	6.84	0.077
Standardized gain	1.16	0.59	1.19	1.39	11.85	0.008
Corrected residual gain	0.27	0.41	0.28	0.42	7.07	0.070

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P^a
QoE/Total						
Variance of:						
Raw gain	0.30	0.12	0.20	0.34	21.91	0.000
Residual gain	0.23	0.10	0.18	0.27	18.30	0.000
Standardized gain	1.35	0.62	0.81	1.30	14.73	0.002
Corrected residual gain	0.24	0.12	0.20	0.28	12.16	0.007
Global Job Satisfaction						
Variance of:						
Raw gain	1.28	1.34	1.07	1.43	2.11	0.550
Residual gain	0.86	1.00	0.82	0.90	0.83	0.843
Standardized gain	1.29	1.13	0.93	1.35	4.04	0.257
Corrected residual gain	0.87	1.04	0.87	0.91	0.71	0.871
JS/Resource Adequacy						
Variance of:						
Raw gain	0.27	0.38	0.60	0.48	14.74	0.002
Residual gain	0.21	0.27	0.44	0.44	16.48	0.001
Standardized gain	1.19	0.62	1.27	1.11	9.69	0.021
Corrected residual gain	0.22	0.33	0.46	0.47	14.61	0.002
JS/Financial Rewards						
Variance of:						
Raw gain	0.61	0.36	0.47	0.78	11.66	0.009
Residual gain	0.47	0.29	0.39	0.48	5.41	0.144
Standardized gain	1.79	0.56	0.80	1.68	34.37	0.000
Corrected residual gain	0.47	0.37	0.45	0.49	1.42	0.701

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P^a
JS/Challenge						
Variance of:						
Raw gain	0.60	0.21	0.44	0.47	18.08	0.000
Residual gain	0.36	0.16	0.29	0.38	13.59	0.004
Standardized gain	1.67	0.53	0.96	1.12	21.93	0.000
Corrected residual gain	0.36	0.17	0.30	0.38	12.35	0.006
JS/Comfort						
Variance of:						
Raw gain	0.36	0.18	0.27	0.34	9.63	0.022
Residual gain	0.23	0.16	0.22	0.29	6.96	0.073
Standardized gain	1.90	0.58	0.87	0.91	28.18	0.000
Corrected residual gain	0.23	0.20	0.24	0.33	5.18	0.159
JS/Coworkers						
Variance of:						
Raw gain	0.53	0.39	0.53	0.91	14.96	0.002
Residual gain	0.34	0.19	0.31	0.36	7.62	0.055
Standardized gain	1.86	1.66	1.67	1.60	0.53	0.911
Corrected residual gain	0.34	0.19	0.32	0.37	7.59	0.055
Ambiguity						
Variance of:						
Raw gain	0.40	0.63	0.35	0.73	15.07	0.002
Residual gain	0.31	0.24	0.25	0.45	10.47	0.015
Standardized gain	1.15	2.06	1.22	1.89	10.09	0.018
Corrected residual gain	0.36	0.24	0.28	0.45	8.60	0.035

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P^a
Effort Expended						
Variance of:						
Raw gain	0.54	0.52	0.79	0.61	4.55	0.208
Residual gain	0.41	0.23	0.59	0.30	18.89	0.000
Standardized gain	1.21	1.00	1.33	1.07	1.96	0.580
Corrected residual gain	0.54	0.35	0.71	0.44	10.13	0.018
Supervisory Rating						
Variance of:						
Raw gain	1.73	2.16	0.14	0.77	24.47	0.000
Residual gain	1.05	1.90	0.09	0.62	27.50	0.000
Standardized gain	1.89	1.04	0.05	2.18	36.92	0.000
Corrected residual gain	1.05	1.90	0.10	0.62	25.80	0.000
Absences[b]						
Variance of:						
Raw gain	9.51	15.56	45.32	10.33	79.30	0.000
Residual gain	0.47	5.32	17.55	10.02	189.28	0.000
Standardized gain	2.32	1.66	1.82	1.78	2.42	0.491
Corrected residual gain	--	--	--	--	--	---
Involvement[b]						
Variance of:						
Raw gain	0.63	0.94	0.93	1.19	7.63	0.054
Residual gain	0.34	0.61	0.72	0.71	13.16	0.004
Standardized gain	1.26	1.23	1.04	1.42	2.22	0.528
Corrected residual gain	--	--	--	--	--	---

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	X^2(3)	P[a]
Intent to Turnover[b]						
Variance of:						
Raw gain	3.53	2.94	1.42	2.24	20.61	0.000
Residual gain	2.68	1.76	0.81	1.27	34.11	0.000
Standardized gain	1.25	1.57	1.44	1.05	3.47	0.324
Corrected residual gain	--	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimates, variance of corrected residual scores can not be computed.

cant instances of heteroscedasticity the smallest variance was found among "clerical and kindred" workers. This is found for measures of working conditions, worker attitudes and worker behaviors alike, and for all four gain scores.

Inspection of the correlations of the variables with themselves over time (i.e., the variables' autocorrelations) in the different occupational groups showed that in 10 of the 15 variables for which heteroscedasticity was found, the highest autocorrelation was observed in the clericals' occupational group. These high correlations accordingly reduce the variance of the gain scores.

To summarize these results on job characteristics:

1. Low income workers typically display more variance over time in work-related measures than higher income workers. The greater variance of attitudinal and behavioral measures among the low income workers is, by and large, restricted to raw gain scores. Differences between the income groups in variance of measures of working conditions are found for both raw gain and standardized gain scores;

2. Occupational group is a strong source of heteroscedasticity of gain scores of work-related measures. Clericals are usually the category displaying the least variance of changes over time.

Personal Demographics. The effects of personal demographic variables (race, sex, age, tenure and education) are meaningfully studied over time only among people who retain these characteristics (or at worst have a constant added to them, for example the workers were all 20 months older by Phase II). Accordingly, the jobs that experienced turnover and replacement were removed from this analysis, since the occupants of such jobs (and their demographics) may have changed in the interim.

A level of a particular variable was retained only if it contained 25 or more cases. The levels of the variables retained are: race (black vs. white), sex, tenure (trichotomized into 1-2 years, 3-5 years and 6+ years at Phase I), education (1-8 years, 9-11 years, 12 years, and 13-15 years), and age (22-29, 30-44 and 45-54 years at Phase I). Variables with two levels were tested by use of a two-tailed F test, in the same manner as income (above). For variables with more than two categories, Bartlett's test for homogeneity of variance was used.

Race-related differences in variance over time were found for six variables, five of which are measures of worker attitudes and behavior. These five variables (global job satisfaction, JS/Coworkers, effort expended, absences and involvement) all display more variance of raw gain scores among black workers than white workers. No significant differences were found for the other gain scores of these variables. Inspection of the components of variance of raw gains revealed that these differences in raw gain score variance were consistently related to differences in the variances of the two races during the first wave of measurement, a difference that was reduced by Phase II. Specifically, there was markedly more variance among black workers than white workers in Phase I (in the winter of 1972-73) than was the case in Phase II (in the fall of 1974). Since raw gain is the only gain score of these four that is sensitive to variance at time one, it follows that raw gain would be the only gain score type to be affected by the differences in variance at Phase I. It is not obvious why differences in the amount of variance in worker attitudes and behavior should exist between the races in 1972-73 or why these differences should be reduced by 1974.

National survey data indicate that for at least one worker attitude, an overall measure of job satisfaction, there was more variance among black workers than white workers in 1969, a difference that is reduced by 1973.[3] It would seem then that the differences in variances between the two races were not unique to the sites studied in the present investigation, but rather were a nationwide phenomenon. The difference between races in variances might then be a function of the particular historical period studied. Although these differences were present in 1969 (and perhaps earlier too) they had diminished by 1973 and 1974. If such within-phase differences in variance between the races do not reappear, then race-related differences of variance of gain scores in worker attitudes and behavior are not likely to be found again, at least for U.S. samples.

One measure of working conditions (QoE/Comfort) displayed significant race-related differences in variance of residual gain ($F_{51,166} = 1.79$, $p < 0.006$, two-tailed) and corrected residual gain ($F_{51,166} = 1.81$, $p < 0.006$, two-tailed). This greater variance over time among black workers for QoE/Comfort is attributable to the greater variance at Phase II found among black workers in their ratings of this facet of their working conditions.

Sex-related differences in variance over time were found to be significant for four variables: QoE/Challenge, QoE/Comfort, QoE/Total and JS/Coworkers. For the three QoE measures of working conditions, there was significantly more variance over time among women. For QoE/Challenge, significantly more variance over time for women was found for variance of raw gains ($F_{94,128} = 1.52$, $p < 0.028$, two-tailed). For QoE/Comfort, women displayed significantly more variance of residual gain ($F_{94,128} = 1.49$, $p < 0.036$, two-tailed) and corrected residual gain ($F_{92,128} = 1.46$, $p < 0.048$, two-tailed). In the case of QoE/Total, women had significantly more variance of standardized gain ($F_{94,128} = 1.62$, $p < 0.010$, two-tailed).

---

[3] R. P. Quinn, Personal communication, January 7, 1975.

The fourth variable that displayed significant sex-related differences in variance over time was an attitudinal variable, JS/Coworkers. Here, it was the men who displayed significantly more variance over time, for residual gain ($F_{133,97} = 1.47$, $p < 0.044$, two-tailed) and corrected residual gain ($F_{133,97} = 1.61$, $p < 0.014$, two-tailed).

A worker's tenure was found to be a source of heteroscedasticity of at least one type of gain score, for 10 of the 17 work-related variables (see Table 12). Heteroscedasticity was found in nine of these variables for raw gain, in seven variables for residual gain and in six variables for standardized and corrected residual gain variance. In most cases it is the middle range of tenure studied (3-5 years at Phase I) that has the most variance over time. This is likely a result of the relatively lower autocorrelations found in this category of tenure, which are the lowest among the three tenure groups studied, for nine of the 10 significantly heteroscedastic variables. Since this is a rather surprising finding, potential confounding effects with tenure were investigated that might explain why there is more variance of change among middle tenure workers. It was found that this tenure group tended to have more blacks, more women and more low income workers than the other categories of tenure. Since it was found above that blacks, women and low income workers displayed more variance over time, these effects may help explain this larger variance over time among the midrange of tenure.

Worker age was found to be a source of heteroscedasticity for at least one gain score for 12 work-related variables. It can be seen in Table 13

Table 12

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P^a
QoE/Resource Adequacy					
Variance of:					
Raw gain	0.86	0.74	0.64	3.27	0.195
Residual gain	0.65	0.45	0.46	4.58	0.101
Standardized gain	1.40	1.39	1.22	1.63	0.596
Corrected residual gain	0.68	0.47	0.50	4.00	0.136
QoE/Financial Rewards					
Variance of:					
Raw gain	0.28	0.57	0.37	12.73	0.002
Residual gain	0.23	0.38	0.33	5.70	0.058
Standardized gain	0.50	1.50	0.73	31.38	0.000
Corrected residual gain	0.36	0.39	0.43	1.00	0.608
QoE/Challenge					
Variance of:					
Raw gain	0.56	0.70	0.26	50.40	0.000
Residual gain	0.50	0.46	0.24	30.06	0.000
Standardized gain	0.78	0.87	0.49	17.67	0.000
Corrected residual gain	0.63	0.56	0.38	12.26	0.002
QoE/Comfort					
Variance of:					
Raw gain	0.24	0.28	0.34	4.11	0.128
Residual gain	0.17	0.23	0.24	3.72	0.155
Standardized gain	0.88	1.03	1.05	1.04	0.594
Corrected residual gain	0.21	0.26	0.27	2.50	0.287
QoE/Total					
Variance of:					
Raw gain	0.22	0.26	0.13	22.42	0.000
Residual gain	0.20	0.18	0.12	14.08	0.001
Standardized gain	0.68	0.94	0.60	8.65	0.013
Corrected residual gain	0.24	0.20	0.15	10.03	0.007

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	p^a
Global Job Satisfaction					
Variance of:					
Raw gain	0.82	1.27	0.83	8.58	0.014
Residual gain	0.59	0.93	0.65	7.03	0.030
Standardized gain	0.98	1.14	0.92	1.94	0.379
Corrected residual gain	0.62	0.96	0.68	6.27	0.044
JS/Resource Adequacy					
Variance of:					
Raw gain	0.45	0.50	0.43	0.91	0.633
Residual gain	0.34	0.34	0.33	0.05	0.977
Standardized gain	1.05	0.96	0.99	0.18	0.911
Corrected residual gain	0.36	0.37	0.36	0.05	0.974
JS/Financial Rewards					
Variance of:					
Raw gain	0.43	0.66	0.33	20.95	0.000
Residual gain	0.34	0.46	0.26	14.18	0.001
Standardized gain	0.81	1.06	0.86	2.39	0.303
Corrected residual gain	0.39	0.50	0.29	12.07	0.002
JS/Challenge					
Variance of:					
Raw gain	0.31	0.46	0.29	9.51	0.009
Residual gain	0.26	0.36	0.24	6.87	0.032
Standardized gain	0.60	1.06	0.76	8.15	0.017
Corrected residual gain	0.27	0.37	0.25	6.33	0.042
JS/Comfort					
Variance of:					
Raw gain	0.19	0.36	0.28	9.49	0.009
Residual gain	0.16	0.26	0.21	5.72	0.057
Standardized gain	0.92	1.13	0.97	1.28	0.528
Corrected residual gain	0.17	0.27	0.23	4.77	0.092

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P^a
JS/Coworkers					
Variance of:					
Raw gain	0.40	0.63	0.51	5.11	0.078
Residual gain	0.25	0.33	0.27	1.99	0.370
Standardized gain	1.32	1.66	1.39	1.70	0.428
Corrected residual gain	0.28	0.34	0.30	0.73	0.695
Ambiguity					
Variance of:					
Raw gain	0.26	0.64	0.36	22.24	0.000
Residual gain	0.23	0.26	0.26	0.45	0.797
Standardized gain	0.80	1.59	1.05	12.48	0.002
Corrected residual gain	0.33	0.27	0.32	1.39	0.500
Effort Expended					
Variance of:					
Raw gain	0.56	0.50	0.53	0.39	0.821
Residual gain	0.50	0.35	0.36	4.60	0.100
Standardized gain	0.97	0.91	0.96	0.14	0.931
Corrected residual gain	0.79	0.58	0.57	3.91	0.142
Supervisory Rating					
Variance of:					
Raw gain	1.27	1.09	1.50	0.90	0.64
Residual gain	1.19	1.03	1.34	0.59	0.74
Standardized gain	0.72	0.78	1.07	1.39	0.50
Corrected residual gain	1.19	1.04	1.34	0.58	0.75
Absences[b]					
Variance of:					
Raw gain	2.40	10.23	19.86	109.10	0.000
Residual gain	1.01	1.11	5.17	124.49	0.000
Standardized gain	1.53	2.22	1.50	6.76	0.034
Corrected residual gain	--	--	--	--	---

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P[a]
Involvement[b]					
Variance of:					
Raw gain	1.03	0.90	0.74	4.47	0.107
Residual gain	0.74	0.50	0.58	3.74	0.154
Standardized gain	1.22	1.22	1.05	1.37	0.503
Corrected residual gain	--	--	--	--	---
Intent to Turnover[b]					
Variance of:					
Raw gain	2.87	3.07	1.39	36.15	0.000
Residual gain	2.58	2.40	1.06	45.36	0.000
Standardized gain	1.27	1.20	1.26	0.10	0.951
Corrected residual gain	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimate, variance of corrected residual scores can not be computed.

Table 13

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P[a]
QoE/Resource Adequacy					
Variance of:					
Raw gain	1.11	0.48	0.51	31.99	0.000
Residual gain	0.69	0.31	0.35	28.16	0.000
Standardized gain	1.69	1.28	0.93	9.82	0.007
Corrected residual gain	0.70	0.33	0.41	22.68	0.000
QoE/Financial Rewards					
Variance of:					
Raw gain	0.40	0.30	0.38	3.65	0.161
Residual gain	0.34	0.26	0.26	2.95	0.228
Standardized gain	0.81	0.65	0.83	2.64	0.267
Corrected residual gain	0.42	0.36	0.32	2.17	0.337
QoE/Challenge					
Variance of:					
Raw gain	0.50	0.38	0.38	3.26	0.196
Residual gain	0.42	0.37	0.25	7.70	0.021
Standardized gain	0.75	0.62	0.60	1.96	0.375
Corrected residual gain	0.54	0.53	0.36	5.58	0.061
QoE/Comfort					
Variance of:					
Raw gain	0.28	0.28	0.30	0.14	0.931
Residual gain	0.25	0.19	0.19	3.73	0.155
Standardized gain	0.90	1.10	0.99	1.59	0.452
Corrected residual gain	0.29	0.21	0.22	5.49	0.064
QoE/Total					
Variance of:					
Raw gain	0.22	0.15	0.15	8.22	0.016
Residual gain	0.19	0.14	0.10	12.56	0.002
Standardized gain	0.88	0.64	0.65	4.98	0.083
Corrected residual gain	0.21	0.18	0.12	9.68	0.008

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	p^a
Global Job Satisfaction					
Variance of:					
Raw gain	1.28	0.68	0.93	16.46	0.000
Residual gain	1.01	0.55	0.54	18.93	0.000
Standardized gain	1.12	0.99	1.27	2.27	0.321
Corrected residual gain	1.04	0.58	0.55	18.74	0.000
JS/Resource Adequacy					
Variance of:					
Raw gain	0.56	0.48	0.48	1.08	0.583
Residual gain	0.42	0.30	0.35	4.44	0.109
Standardized gain	1.11	1.35	0.90	5.84	0.054
Corrected residual gain	0.44	0.31	0.40	5.60	0.061
JS/Financial Rewards					
Variance of:					
Raw gain	0.37	0.46	0.48	3.01	0.222
Residual gain	0.30	0.34	0.27	2.17	0.339
Standardized gain	0.75	1.14	1.29	10.27	0.006
Corrected residual gain	0.35	0.36	0.28	2.52	0.283
JS/Challenge					
Variance of:					
Raw gain	0.43	0.27	0.39	10.30	0.006
Residual gain	0.37	0.22	0.26	12.01	0.003
Standardized gain	0.76	0.75	1.22	10.08	0.007
Corrected residual gain	0.38	0.23	0.26	12.15	0.002
JS/Comfort					
Variance of:					
Raw gain	0.25	0.24	0.43	13.99	0.001
Residual gain	0.20	0.18	0.23	2.44	0.296
Standardized gain	1.13	0.98	1.37	3.83	0.147
Corrected residual gain	0.21	0.20	0.24	1.46	0.483

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	X^2(2)	P[a]
JS/Coworkers					
Variance of:					
Raw gain	0.56	0.35	0.75	22.11	0.000
Residual gain	0.24	0.21	0.28	2.36	0.307
Standardized gain	1.75	1.08	1.74	12.34	0.002
Corrected residual gain	0.25	0.28	0.28	0.64	0.705
Ambiguity					
Variance of:					
Raw gain	0.41	0.32	0.58	11.77	0.003
Resdiual gain	0.24	0.25	0.27	0.38	0.825
Standardized gain	1.34	0.90	1.66	14.30	0.001
Corrected residual gain	0.27	0.33	0.28	2.33	0.312
Effort Expended					
Variance of:					
Raw gain	0.49	0.49	0.63	2.55	0.279
Residual gain	0.43	0.36	0.39	1.49	0.475
Standardized gain	0.94	0.90	0.98	0.26	0.878
Corrected residual gain	0.71	0.60	0.60	1.20	0.550
Supervisory Rating					
Variance of:					
Raw gain	1.29	1.86	1.45	0.83	0.662
Residual gain	1.29	0.50	1.15	4.79	0.091
Standardized gain	0.76	2.10	1.75	8.05	0.018
Corrected residual gain	1.29	0.50	1.15	4.83	0.089
Absences[b]					
Variance of:					
Raw gain	11.75	15.63	22.97	14.34	0.001
Residual gain	1.94	3.92	6.78	47.92	0.000
Standardized gain	1.37	1.63	1.58	1.24	0.539
Corrected residual gain	--	--	--	--	---

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P[a]
Involvement[b]					
Variance of:					
Raw gain	0.83	0.82	0.71	0.90	0.640
Residual gain	0.68	0.56	0.45	3.83	0.150
Standardized gain	1.08	1.26	1.15	0.97	0.610
Corrected residual gain	--	--	--	--	---
Intent to Turnover[b]					
Variance of:					
Raw gain	2.62	1.80	1.20	18.46	0.000
Residual gain	2.35	1.21	0.62	52.65	0.000
Standardized gain	1.06	1.50	1.34	4.70	0.096
Corrected residual gain	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimate, variance of corrected residual scores can not be computed.

that raw gain and residual gain are change score types most often affected by age and so display significant heteroscedasticity for nine and seven of the 12 variables, respectively. Less often affected are standardized gain (six variables) and corrected residual gain (four variables). The younger workers in the sample (22-29 at Phase I) tended to have the most variance over time and the middle age group (30-44 at Phase I) tended to have the least variance over time, among the instances of significant heteroscedasticity.

Educational level was a source of heteroscedasticity for at least one gain score of 11 of the work-related variables (see Table 14). Heteroscedasticity was most strongly found among the six measures of working conditions, five of which displayed educational-level-related heteroscedasticity. Raw gain and residual gain score variances were particularly affected by educational level, displaying heteroscedasticity for 10 and nine variables, respectively. Standardized gain and corrected residual gain were affected, respectively, for five and four variables.

Workers with the highest level of education studied here (some college but not completed) were typically the most stable category in their ratings of working conditions, while those with the least education (1-8 years) tended to have the most variance over time in their ratings of working conditions. A somewhat different pattern of results can be seen among those measures of worker attitudes and behaviors that displayed significant heteroscedasticity. Workers with some high school (but who had not completed it) tended to display the least variance over time among the significantly heteroscedastic attitude and behavior variables. No one education category typically displayed the highest variance of gain scores

Table 14

Homogeneity of Variance tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13-15 Years	$X^2(3)$	P^a
QoE/Resource Adequacy						
Variance of:						
Raw gain	0.88	0.85	0.69	0.52	7.85	0.049
Residual gain	0.57	0.52	0.53	0.33	8.97	0.030
Standardized gain	1.32	1.75	1.06	1.04	10.33	0.016
Corrected residual gain	0.60	0.53	0.60	0.37	7.67	0.053
QoE/Financial Rewards						
Variance of:						
Raw gain	0.52	0.26	0.41	0.53	15.75	0.001
Residual gain	0.46	0.21	0.31	0.43	17.76	0.001
Standardized gain	0.85	0.52	0.97	1.11	16.82	0.001
Corrected residual gain	0.56	0.32	0.36	0.48	8.60	0.035
QoE/Challenge						
Variance of:						
Raw gain	0.37	0.37	0.54	0.32	10.39	0.016
Residual gain	0.27	0.30	0.41	0.31	6.53	0.089
Standardized gain	0.60	0.69	0.75	0.49	6.42	0.093
Corrected residual gain	0.38	0.40	0.53	0.49	4.00	0.262
QoE/Comfort						
Variance of:						
Raw gain	0.43	0.36	0.31	0.30	3.64	0.303
Residual gain	0.28	0.21	0.25	0.20	2.21	0.530
Standardized gain	1.42	1.05	1.05	1.00	2.96	0.399
Corrected residual gain	0.29	0.27	0.29	0.24	1.43	0.699

Table 10 continued

Homogeneity of Variance Tests for Variables with Significant Differences in Variances over Time between the Above and Below Median Income Groups

	df[a]	Variance of Raw Gain	F[b]	p[c]	Variance of Standardized Gain	F[b]	p[c]	Variance of Residual Gain	F[b]	p[c]	Variance of Corrected Residual Gain	F[b]	p[c]
JS/Resource Adequacy													
Low income	102	0.38	1.45	0.070	0.85	1.49	0.050	0.29	1.37	0.126	0.33	1.26	0.260
High income	90	0.55			1.27			0.40			0.42		
JS/Financial Rewards													
Low income	101	0.70	1.52	0.042	1.32	1.25	0.278	0.47	1.27	0.250	0.49	1.22	0.338
High income	90	0.46			1.06			0.37			0.40		
Absences													
Low income	101	28.59	1.63	0.018	1.97	1.35	0.142	13.26	2.38	0.000	_d		
High income	91	17.56			1.46			5.58					
Intent to Turnover													
Low income	101	2.81	1.61	0.022	1.36	1.06	0.674	1.75	1.25	0.276	_d		
High income	90	1.75			1.28			1.40					

[a] degrees of freedom change due to missing data

[b] F's always are computed by dividing the larger variance by the smaller

[c] Probabilities are exact, two-tailed probabilities rounded to the nearest thousandth

[d] No reliability estimates available, so corrected residual scores can not be computed.

Table 11

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P^a
QoE/Resource Adequacy						
Variance of:						
Raw gain	0.82	0.54	0.91	0.61	6.55	0.088
Residual gain	0.60	0.32	0.58	0.50	7.19	0.066
Standardized gain	1.89	0.68	1.48	1.35	17.03	0.001
Corrected residual gain	0.60	0.44	0.59	0.53	2.15	0.543
QoE/Financial Rewards						
Variance of:						
Raw gain	1.47	0.31	0.42	0.52	55.72	0.000
Residual gain	0.92	0.20	0.37	0.44	39.55	0.000
Standardized gain	2.08	1.37	0.80	1.05	22.05	0.000
Corrected residual gain	0.92	0.21	0.47	0.49	33.04	0.000
QoE/Challenge						
Variance of:						
Raw gain	0.63	0.35	0.65	0.85	13.18	0.004
Residual gain	0.51	0.20	0.44	0.58	18.55	0.000
Standardized gain	1.29	0.74	1.10	1.25	6.00	0.112
Corrected residual gain	0.54	0.26	0.49	0.62	12.61	0.006
QoE/Comfort						
Variance of:						
Raw gain	0.37	0.28	0.40	0.50	5.42	0.144
Residual gain	0.24	0.28	0.25	0.40	6.84	0.077
Standardized gain	1.16	0.59	1.19	1.39	11.85	0.008
Corrected residual gain	0.27	0.41	0.28	0.42	7.07	0.070

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P[a]
QoE/Total						
Variance of:						
Raw gain	0.30	0.12	0.20	0.34	21.91	0.000
Residual gain	0.23	0.10	0.18	0.27	18.30	0.000
Standardized gain	1.35	0.62	0.81	1.30	14.73	0.002
Corrected residual gain	0.24	0.12	0.20	0.28	12.16	0.007
Global Job Satisfaction						
Variance of:						
Raw gain	1.28	1.34	1.07	1.43	2.11	0.550
Residual gain	0.86	1.00	0.82	0.90	0.83	0.843
Standardized gain	1.29	1.13	0.93	1.35	4.04	0.257
Corrected residual gain	0.87	1.04	0.87	0.91	0.71	0.871
JS/Resource Adequacy						
Variance of:						
Raw gain	0.27	0.38	0.60	0.48	14.74	0.002
Residual gain	0.21	0.27	0.44	0.44	16.48	0.001
Standardized gain	1.19	0.62	1.27	1.11	9.69	0.021
Corrected residual gain	0.22	0.33	0.46	0.47	14.61	0.002
JS/Financial Rewards						
Variance of:						
Raw gain	0.61	0.36	0.47	0.78	11.66	0.009
Residual gain	0.47	0.29	0.39	0.48	5.41	0.144
Standardized gain	1.79	0.56	0.80	1.68	34.37	0.000
Corrected residual gain	0.47	0.37	0.45	0.49	1.42	0.701

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	p^a
JS/Challenge						
Variance of:						
Raw gain	0.60	0.21	0.44	0.47	18.08	0.000
Residual gain	0.36	0.16	0.29	0.38	13.59	0.004
Standardized gain	1.67	0.53	0.96	1.12	21.93	0.000
Corrected residual gain	0.36	0.17	0.30	0.38	12.35	0.006
JS/Comfort						
Variance of:						
Raw gain	0.36	0.18	0.27	0.34	9.63	0.022
Residual gain	0.23	0.16	0.22	0.29	6.96	0.073
Standardized gain	1.90	0.58	0.87	0.91	28.18	0.000
Corrected residual gain	0.23	0.20	0.24	0.33	5.18	0.159
JS/Coworkers						
Variance of:						
Raw gain	0.53	0.39	0.53	0.91	14.96	0.002
Residual gain	0.34	0.19	0.31	0.36	7.62	0.055
Standardized gain	1.86	1.66	1.67	1.60	0.53	0.911
Corrected residual gain	0.34	0.19	0.32	0.37	7.59	0.055
Ambiguity						
Variance of:						
Raw gain	0.40	0.63	0.35	0.73	15.07	0.002
Residual gain	0.31	0.24	0.25	0.45	10.47	0.015
Standardized gain	1.15	2.06	1.22	1.89	10.09	0.018
Corrected residual gain	0.36	0.24	0.28	0.45	8.60	0.035

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	p^a
Effort Expended						
Variance of:						
Raw gain	0.54	0.52	0.79	0.61	4.55	0.208
Residual gain	0.41	0.23	0.59	0.30	18.89	0.000
Standardized gain	1.21	1.00	1.33	1.07	1.96	0.580
Corrected residual gain	0.54	0.35	0.71	0.44	10.13	0.018
Supervisory Rating						
Variance of:						
Raw gain	1.73	2.16	0.14	0.77	24.47	0.000
Residual gain	1.05	1.90	0.09	0.62	27.50	0.000
Standardized gain	1.89	1.04	0.05	2.18	36.92	0.000
Corrected residual gain	1.05	1.90	0.10	0.62	25.80	0.000
Absences[b]						
Variance of:						
Raw gain	9.51	15.56	45.32	10.33	79.30	0.000
Residual gain	0.47	5.32	17.55	10.02	189.28	0.000
Standardized gain	2.32	1.66	1.82	1.78	2.42	0.491
Corrected residual gain	--	--	--	--	--	---
Involvement[b]						
Variance of:						
Raw gain	0.63	0.94	0.93	1.19	7.63	0.054
Residual gain	0.34	0.61	0.72	0.71	13.16	0.004
Standardized gain	1.26	1.23	1.04	1.42	2.22	0.528
Corrected residual gain	--	--	--	--	--	---

Table 11 continued

Homogeneity of Variance Tests for Occupational Group

Variable	Professionals	Clericals	Operatives	Service workers	$X^2(3)$	P[a]
Intent to Turnover[b]						
Variance of:						
Raw gain	3.53	2.94	1.42	2.24	20.61	0.000
Residual gain	2.68	1.76	0.81	1.27	34.11	0.000
Standardized gain	1.25	1.57	1.44	1.05	3.47	0.324
Corrected residual gain	--	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimates, variance of corrected residual scores can not be computed.

cant instances of heteroscedasticity the smallest variance was found among "clerical and kindred" workers. This is found for measures of working conditions, worker attitudes and worker behaviors alike, and for all four gain scores.

Inspection of the correlations of the variables with themselves over time (i.e., the variables' autocorrelations) in the different occupational groups showed that in 10 of the 15 variables for which heteroscedasticity was found, the highest autocorrelation was observed in the clericals' occupational group. These high correlations accordingly reduce the variance of the gain scores.

To summarize these results on job characteristics:

1. Low income workers typically display more variance over time in work-related measures than higher income workers. The greater variance of attitudinal and behavioral measures among the low income workers is, by and large, restricted to raw gain scores. Differences between the income groups in variance of measures of working conditions are found for both raw gain and standardized gain scores;

2. Occupational group is a strong source of heteroscedasticity of gain scores of work-related measures. Clericals are usually the category displaying the least variance of changes over time.

Personal Demographics. The effects of personal demographic variables (race, sex, age, tenure and education) are meaningfully studied over time only among people who retain these characteristics (or at worst have a constant added to them, for example the workers were all 20 months older by Phase II). Accordingly, the jobs that experienced turnover and replacement were removed from this analysis, since the occupants of such jobs (and their demographics) may have changed in the interim.

A level of a particular variable was retained only if it contained 25 or more cases. The levels of the variables retained are: race (black vs. white), sex, tenure (trichotomized into 1-2 years, 3-5 years and 6+ years at Phase I), education (1-8 years, 9-11 years, 12 years, and 13-15 years), and age (22-29, 30-44 and 45-54 years at Phase I). Variables with two levels were tested by use of a two-tailed F test, in the same manner as income (above). For variables with more than two categories, Bartlett's test for homogeneity of variance was used.

Race-related differences in variance over time were found for six variables, five of which are measures of worker attitudes and behavior. These five variables (global job satisfaction, JS/Coworkers, effort expended, absences and involvement) all display more variance of raw gain scores among black workers than white workers. No significant differences were found for the other gain scores of these variables. Inspection of the components of variance of raw gains revealed that these differences in raw gain score variance were consistently related to differences in the variances of the two races during the first wave of measurement, a difference that was reduced by Phase II. Specifically, there was markedly more variance among black workers than white workers in Phase I (in the winter of 1972-73) than was the case in Phase II (in the fall of 1974). Since raw gain is the only gain score of these four that is sensitive to variance at time one, it follows that raw gain would be the only gain score type to be affected by the differences in variance at Phase I. It is not obvious why differences in the amount of variance in worker attitudes and behavior should exist between the races in 1972-73 or why these differences should be reduced by 1974.

National survey data indicate that for at least one worker attitude, an overall measure of job satisfaction, there was more variance among black workers than white workers in 1969, a difference that is reduced by 1973.[3] It would seem then that the differences in variances between the two races were not unique to the sites studied in the present investigation, but rather were a nationwide phenomenon. The difference between races in variances might then be a function of the particular historical period studied. Although these differences were present in 1969 (and perhaps earlier too) they had diminished by 1973 and 1974. If such within-phase differences in variance between the races do not reappear, then race-related differences of variance of gain scores in worker attitudes and behavior are not likely to be found again, at least for U.S. samples.

One measure of working conditions (QoE/Comfort) displayed significant race-related differences in variance of residual gain ($F_{51,166} = 1.79$, $p < 0.006$, two-tailed) and corrected residual gain ($F_{51,166} = 1.81$, $p < 0.006$, two-tailed). This greater variance over time among black workers for QoE/Comfort is attributable to the greater variance at Phase II found among black workers in their ratings of this facet of their working conditions.

Sex-related differences in variance over time were found to be significant for four variables: QoE/Challenge, QoE/Comfort, QoE/Total and JS/Coworkers. For the three QoE measures of working conditions, there was significantly more variance over time among women. For QoE/Challenge, significantly more variance over time for women was found for variance of raw gains ($F_{94,128} = 1.52$, $p < 0.028$, two-tailed). For QoE/Comfort, women displayed significantly more variance of residual gain ($F_{94,128} = 1.49$, $p < 0.036$, two-tailed) and corrected residual gain ($F_{92,128} = 1.46$, $p < 0.048$, two-tailed). In the case of QoE/Total, women had significantly more variance of standardized gain ($F_{94,128} = 1.62$, $p < 0.010$, two-tailed).

---

[3] R. P. Quinn, Personal communication, January 7, 1975.

The fourth variable that displayed significant sex-related differences in variance over time was an attitudinal variable, JS/Coworkers. Here, it was the men who displayed significantly more variance over time, for residual gain ($F_{133,97} = 1.47$, $p < 0.044$, two-tailed) and corrected residual gain ($F_{133,97} = 1.61$, $p < 0.014$, two-tailed).

A worker's tenure was found to be a source of heteroscedasticity of at least one type of gain score, for 10 of the 17 work-related variables (see Table 12). Heteroscedasticity was found in nine of these variables for raw gain, in seven variables for residual gain and in six variables for standardized and corrected residual gain variance. In most cases it is the middle range of tenure studied (3-5 years at Phase I) that has the most variance over time. This is likely a result of the relatively lower autocorrelations found in this category of tenure, which are the lowest among the three tenure groups studied, for nine of the 10 significantly heteroscedastic variables. Since this is a rather surprising finding, potential confounding effects with tenure were investigated that might explain why there is more variance of change among middle tenure workers. It was found that this tenure group tended to have more blacks, more women and more low income workers than the other categories of tenure. Since it was found above that blacks, women and low income workers displayed more variance over time, these effects may help explain this larger variance over time among the midrange of tenure.

Worker age was found to be a source of heteroscedasticity for at least one gain score for 12 work-related variables. It can be seen in Table 13

Table 12

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P^a
QoE/Resource Adequacy					
Variance of:					
Raw gain	0.86	0.74	0.64	3.27	0.195
Residual gain	0.65	0.45	0.46	4.58	0.101
Standardized gain	1.40	1.39	1.22	1.63	0.596
Corrected residual gain	0.68	0.47	0.50	4.00	0.136
QoE/Financial Rewards					
Variance of:					
Raw gain	0.28	0.57	0.37	12.73	0.002
Residual gain	0.23	0.38	0.33	5.70	0.058
Standardized gain	0.50	1.50	0.73	31.38	0.000
Corrected residual gain	0.36	0.39	0.43	1.00	0.608
QoE/Challenge					
Variance of:					
Raw gain	0.56	0.70	0.26	50.40	0.000
Residual gain	0.50	0.46	0.24	30.06	0.000
Standardized gain	0.78	0.87	0.49	17.67	0.000
Corrected residual gain	0.63	0.56	0.38	12.26	0.002
QoE/Comfort					
Variance of:					
Raw gain	0.24	0.28	0.34	4.11	0.128
Residual gain	0.17	0.23	0.24	3.72	0.155
Standardized gain	0.88	1.03	1.05	1.04	0.594
Corrected residual gain	0.21	0.26	0.27	2.50	0.287
QoE/Total					
Variance of:					
Raw gain	0.22	0.26	0.13	22.42	0.000
Residual gain	0.20	0.18	0.12	14.08	0.001
Standardized gain	0.68	0.94	0.60	8.65	0.013
Corrected residual gain	0.24	0.20	0.15	10.03	0.007

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	p^a
Global Job Satisfaction					
Variance of:					
Raw gain	0.82	1.27	0.83	8.58	0.014
Residual gain	0.59	0.93	0.65	7.03	0.030
Standardized gain	0.98	1.14	0.92	1.94	0.379
Corrected residual gain	0.62	0.96	0.68	6.27	0.044
JS/Resource Adequacy					
Variance of:					
Raw gain	0.45	0.50	0.43	0.91	0.633
Residual gain	0.34	0.34	0.33	0.05	0.977
Standardized gain	1.05	0.96	0.99	0.18	0.911
Corrected residual gain	0.36	0.37	0.36	0.05	0.974
JS/Financial Rewards					
Variance of:					
Raw gain	0.43	0.66	0.33	20.95	0.000
Residual gain	0.34	0.46	0.26	14.18	0.001
Standardized gain	0.81	1.06	0.86	2.39	0.303
Corrected residual gain	0.39	0.50	0.29	12.07	0.002
JS/Challenge					
Variance of:					
Raw gain	0.31	0.46	0.29	9.51	0.009
Residual gain	0.26	0.36	0.24	6.87	0.032
Standardized gain	0.60	1.06	0.76	8.15	0.017
Corrected residual gain	0.27	0.37	0.25	6.33	0.042
JS/Comfort					
Variance of:					
Raw gain	0.19	0.36	0.28	9.49	0.009
Residual gain	0.16	0.26	0.21	5.72	0.057
Standardized gain	0.92	1.13	0.97	1.28	0.528
Corrected residual gain	0.17	0.27	0.23	4.77	0.092

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P^a
JS/Coworkers					
Variance of:					
Raw gain	0.40	0.63	0.51	5.11	0.078
Residual gain	0.25	0.33	0.27	1.99	0.370
Standardized gain	1.32	1.66	1.39	1.70	0.428
Corrected residual gain	0.28	0.34	0.30	0.73	0.695
Ambiguity					
Variance of:					
Raw gain	0.26	0.64	0.36	22.24	0.000
Residual gain	0.23	0.26	0.26	0.45	0.797
Standardized gain	0.80	1.59	1.05	12.48	0.002
Corrected residual gain	0.33	0.27	0.32	1.39	0.500
Effort Expended					
Variance of:					
Raw gain	0.56	0.50	0.53	0.39	0.821
Residual gain	0.50	0.35	0.36	4.60	0.100
Standardized gain	0.97	0.91	0.96	0.14	0.931
Corrected residual gain	0.79	0.58	0.57	3.91	0.142
Supervisory Rating					
Variance of:					
Raw gain	1.27	1.09	1.50	0.90	0.64
Residual gain	1.19	1.03	1.34	0.59	0.74
Standardized gain	0.72	0.78	1.07	1.39	0.50
Corrected residual gain	1.19	1.04	1.34	0.58	0.75
Absences[b]					
Variance of:					
Raw gain	2.40	10.23	19.86	109.10	0.000
Residual gain	1.01	1.11	5.17	124.49	0.000
Standardized gain	1.53	2.22	1.50	6.76	0.034
Corrected residual gain	--	--	--	--	---

Table 12 continued

Homogeneity of Variance Tests for Level of Tenure

Variable	1 - 2 Years	3 - 5 Years	6 - + Years	$X^2(2)$	P[a]
Involvement[b]					
Variance of:					
Raw gain	1.03	0.90	0.74	4.47	0.107
Residual gain	0.74	0.50	0.58	3.74	0.154
Standardized gain	1.22	1.22	1.05	1.37	0.503
Corrected residual gain	--	--	--	--	---
Intent to Turnover[b]					
Variance of:					
Raw gain	2.87	3.07	1.39	36.15	0.000
Residual gain	2.58	2.40	1.06	45.36	0.000
Standardized gain	1.27	1.20	1.26	0.10	0.951
Corrected residual gain	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimate, variance of corrected residual scores can not be computed.

Table 13

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P^a
QoE/Resource Adequacy					
Variance of:					
Raw gain	1.11	0.48	0.51	31.99	0.000
Residual gain	0.69	0.31	0.35	28.16	0.000
Standardized gain	1.69	1.28	0.93	9.82	0.007
Corrected residual gain	0.70	0.33	0.41	22.68	0.000
QoE/Financial Rewards					
Variance of:					
Raw gain	0.40	0.30	0.38	3.65	0.161
Residual gain	0.34	0.26	0.26	2.95	0.228
Standardized gain	0.81	0.65	0.83	2.64	0.267
Corrected residual gain	0.42	0.36	0.32	2.17	0.337
QoE/Challenge					
Variance of:					
Raw gain	0.50	0.38	0.38	3.26	0.196
Residual gain	0.42	0.37	0.25	7.70	0.021
Standardized gain	0.75	0.62	0.60	1.96	0.375
Corrected residual gain	0.54	0.53	0.36	5.58	0.061
QoE/Comfort					
Variance of:					
Raw gain	0.28	0.28	0.30	0.14	0.931
Residual gain	0.25	0.19	0.19	3.73	0.155
Standardized gain	0.90	1.10	0.99	1.59	0.452
Corrected residual gain	0.29	0.21	0.22	5.49	0.064
QoE/Total					
Variance of:					
Raw gain	0.22	0.15	0.15	8.22	0.016
Residual gain	0.19	0.14	0.10	12.56	0.002
Standardized gain	0.88	0.64	0.65	4.98	0.083
Corrected residual gain	0.21	0.18	0.12	9.68	0.008

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P[a]
Global Job Satisfaction					
Variance of:					
Raw gain	1.28	0.68	0.93	16.46	0.000
Residual gain	1.01	0.55	0.54	18.93	0.000
Standardized gain	1.12	0.99	1.27	2.27	0.321
Corrected residual gain	1.04	0.58	0.55	18.74	0.000
JS/Resource Adequacy					
Variance of:					
Raw gain	0.56	0.48	0.48	1.08	0.583
Residual gain	0.42	0.30	0.35	4.44	0.109
Standardized gain	1.11	1.35	0.90	5.84	0.054
Corrected residual gain	0.44	0.31	0.40	5.60	0.061
JS/Financial Rewards					
Variance of:					
Raw gain	0.37	0.46	0.48	3.01	0.222
Residual gain	0.30	0.34	0.27	2.17	0.339
Standardized gain	0.75	1.14	1.29	10.27	0.006
Corrected residual gain	0.35	0.36	0.28	2.52	0.283
JS/Challenge					
Variance of:					
Raw gain	0.43	0.27	0.39	10.30	0.006
Residual gain	0.37	0.22	0.26	12.01	0.003
Standardized gain	0.76	0.75	1.22	10.08	0.007
Corrected residual gain	0.38	0.23	0.26	12.15	0.002
JS/Comfort					
Variance of:					
Raw gain	0.25	0.24	0.43	13.99	0.001
Residual gain	0.20	0.18	0.23	2.44	0.296
Standardized gain	1.13	0.98	1.37	3.83	0.147
Corrected residual gain	0.21	0.20	0.24	1.46	0.483

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P^a
JS/Coworkers					
Variance of:					
Raw gain	0.55	0.35	0.75	22.11	0.000
Residual gain	0.24	0.21	0.28	2.36	0.307
Standardized gain	1.75	1.08	1.74	12.34	0.002
Corrected residual gain	0.25	0.28	0.28	0.64	0.705
Ambiguity					
Variance of:					
Raw gain	0.41	0.32	0.58	11.77	0.003
Resdiual gain	0.24	0.25	0.27	0.38	0.825
Standardized gain	1.34	0.90	1.66	14.30	0.001
Corrected residual gain	0.27	0.33	0.28	2.33	0.312
Effort Expended					
Variance of:					
Raw gain	0.49	0.49	0.63	2.55	0.279
Residual gain	0.43	0.36	0.39	1.49	0.475
Standardized gain	0.94	0.90	0.98	0.26	0.878
Corrected residual gain	0.71	0.60	0.60	1.20	0.550
Supervisory Rating					
Variance of:					
Raw gain	1.29	1.86	1.45	0.83	0.662
Residual gain	1.29	0.50	1.15	4.79	0.091
Standardized gain	0.76	2.10	1.75	8.05	0.018
Corrected residual gain	1.29	0.50	1.15	4.83	0.089
Absences[b]					
Variance of:					
Raw gain	11.75	15.63	22.97	14.34	0.001
Residual gain	1.94	3.92	6.78	47.92	0.000
Standardized gain	1.37	1.63	1.58	1.24	0.539
Corrected residual gain	--	--	--	--	---

Table 13 continued

Homogeneity of Variance Tests for Age Level

Variable	22-29 Years	30-44 Years	45-54 Years	$X^2(2)$	P^a
Involvement[b]					
Variance of:					
Raw gain	0.83	0.82	0.71	0.90	0.640
Residual gain	0.68	0.56	0.45	3.83	0.150
Standardized gain	1.08	1.26	1.15	0.97	0.610
Corrected residual gain	--	--	--	--	---
Intent to Turnover[b]					
Variance of:					
Raw gain	2.62	1.80	1.20	18.46	0.000
Residual gain	2.35	1.21	0.62	52.65	0.000
Standardized gain	1.06	1.50	1.34	4.70	0.096
Corrected residual gain	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimate, variance of corrected residual scores can not be computed.

that raw gain and residual gain are change score types most often affected by age and so display significant heteroscedasticity for nine and seven of the 12 variables, respectively. Less often affected are standardized gain (six variables) and corrected residual gain (four variables). The younger workers in the sample (22-29 at Phase I) tended to have the most variance over time and the middle age group (30-44 at Phase I) tended to have the least variance over time, among the instances of significant heteroscedasticity.

Educational level was a source of heteroscedasticity for at least one gain score of 11 of the work-related variables (see Table 14). Heteroscedasticity was most strongly found among the six measures of working conditions, five of which displayed educational-level-related heteroscedasticity. Raw gain and residual gain score variances were particularly affected by educational level, displaying heteroscedasticity for 10 and nine variables, respectively. Standardized gain and corrected residual gain were affected, respectively, for five and four variables.

Workers with the highest level of education studied here (some college but not completed) were typically the most stable category in their ratings of working conditions, while those with the least education (1-8 years) tended to have the most variance over time in their ratings of working conditions. A somewhat different pattern of results can be seen among those measures of worker attitudes and behaviors that displayed significant heteroscedasticity. Workers with some high school (but who had not completed it) tended to display the least variance over time among the significantly heteroscedastic attitude and behavior variables. No one education category typically displayed the highest variance of gain scores

Table 14

Homogeneity of Variance tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13–15 Years	$X^2(3)$	P^a
QoE/Resource Adequacy						
Variance of:						
Raw gain	0.88	0.85	0.69	0.52	7.85	0.049
Residual gain	0.57	0.52	0.53	0.33	8.97	0.030
Standardized gain	1.32	1.75	1.06	1.04	10.33	0.016
Corrected residual gain	0.60	0.53	0.60	0.37	7.67	0.053
QoE/Financial Rewards						
Variance of:						
Raw gain	0.52	0.26	0.41	0.53	15.75	0.001
Residual gain	0.46	0.21	0.31	0.43	17.76	0.001
Standardized gain	0.85	0.52	0.97	1.11	16.82	0.001
Corrected residual gain	0.56	0.32	0.36	0.48	8.60	0.035
QoE/Challenge						
Variance of:						
Raw gain	0.37	0.37	0.54	0.32	10.39	0.016
Residual gain	0.27	0.30	0.41	0.31	6.53	0.089
Standardized gain	0.60	0.69	0.75	0.49	6.42	0.093
Corrected residual gain	0.38	0.40	0.53	0.49	4.00	0.262
QoE/Comfort						
Variance of:						
Raw gain	0.43	0.36	0.31	0.30	3.64	0.303
Residual gain	0.28	0.21	0.25	0.20	2.21	0.530
Standardized gain	1.42	1.05	1.05	1.00	2.96	0.399
Corrected residual gain	0.29	0.27	0.29	0.24	1.43	0.699

Table 14 continued

Homogeneity of Variance Tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13-15 Years	$X^2(3)$	P^a
QoE/Total						
Variance of:						
Raw gain	0.19	0.16	0.22	0.13	9.67	0.022
Residual gain	0.15	0.13	0.19	0.12	7.13	0.068
Standardized gain	0.83	0.82	0.78	0.51	7.85	0.049
Corrected residual gain	0.17	0.15	0.22	0.17	5.23	0.156
Global Job Satisfaction						
Variance of:						
Raw gain	0.65	0.89	0.96	1.23	8.71	0.034
Residual gain	0.40	0.61	0.80	0.90	14.52	0.002
Standardized gain	0.80	1.16	0.94	1.13	4.04	0.257
Corrected residual gain	0.43	0.63	0.81	0.93	13.72	0.003
JS/Resource Adequacy						
Variance of:						
Raw gain	0.51	0.45	0.46	0.53	0.89	0.827
Residual gain	0.32	0.39	0.33	0.30	1.88	0.599
Standardized gain	0.97	0.91	0.98	1.22	2.89	0.408
Corrected residual gain	0.35	0.43	0.36	0.31	2.77	0.428
JS/Financial Rewards						
Variance of:						
Raw gain	0.40	0.33	0.41	0.48	3.91	0.272
Residual gain	0.30	0.26	0.32	0.40	5.59	0.133
Standardized gain	1.00	0.77	0.74	1.04	3.39	0.146
Corrected residual gain	0.33	0.30	0.38	0.44	4.19	0.241

Table 14 continued

Homogeneity of Variance Tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13-15 Years	$X^2(3)$	P^a
JS/Challenge						
Variance of:						
Raw gain	0.36	0.32	0.34	0.40	1.64	0.651
Residual gain	0.26	0.26	0.27	0.35	3.29	0.348
Standardized gain	1.09	0.78	0.73	0.94	5.20	0.158
Corrected residual gain	0.27	0.26	0.28	0.35	3.15	0.369
JS/Comfort						
Variance of:						
Raw gain	0.33	0.26	0.30	0.26	1.86	0.603
Residual gain	0.19	0.17	0.26	0.19	6.90	0.075
Standardized gain	1.21	0.84	0.92	1.15	4.48	0.214
Corrected residual gain	0.20	0.19	0.28	0.21	7.29	0.063
JS/Coworkers						
Variance of:						
Raw gain	0.59	0.38	0.66	0.51	10.70	0.014
Residual gain	0.23	0.22	0.38	0.22	16.52	0.001
Standardized gain	2.29	1.21	1.55	1.28	10.64	0.014
Corrected residual gain	0.23	0.27	0.40	0.25	12.22	0.007
Ambiguity						
Variance of:						
Raw gain	0.59	0.39	0.48	0.26	17.35	0.001
Residual gain	0.22	0.30	0.31	0.16	14.45	0.002
Standardized gain	2.06	1.26	1.29	0.58	34.33	0.000
Corrected residual gain	0.22	0.33	0.34	0.28	4.33	0.228

Table 14 continued

Homogeneity of Variance Tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13-15 Years	$X^2(3)$	P[a]
Effort Expended						
Variance of:						
Raw gain	0.70	0.63	0.48	0.61	4.43	0.219
Residual gain	0.52	0.40	0.35	0.39	4.29	0.231
Standardized gain	1.16	1.12	0.97	1.12	1.25	0.741
Corrected residual gain	0.70	0.56	0.55	0.54	1.87	0.600
Supervisory Rating						
Variance of:						
Raw gain	1.00	1.25	2.50	1.08	5.15	0.161
Residual gain	0.33	1.13	2.16	1.07	7.40	0.060
Standardized gain	2.00	1.53	0.89	0.72	5.66	0.129
Corrected residual gain	0.33	1.13	2.16	1.68	7.41	0.060
Absences[b]						
Variance of:						
Raw gain	65.82	31.86	5.16	5.28	264.37	0.000
Residual gain	28.79	9.17	2.91	0.96	292.34	0.000
Standardized gain	1.78	1.80	1.63	1.96	1.27	0.735
Corrected residual gain	--	--	--	--	--	---
Involvement[b]						
Variance of:						
Raw gain	1.29	0.88	0.91	0.61	12.02	0.007
Residual gain	0.94	0.71	0.52	0.50	12.60	0.006
Standardized gain	1.61	1.14	1.16	1.01	4.81	0.186
Corrected residual gain	--	--	--	--	--	---

Table 14 continued

Homogeneity of Variance Tests for Educational Level

Variable	1 - 8 Years	9 -11 Years	12 Years	13-15 Years	$X^2(3)$	p[a]
Intent to Turnover						
Variance of:						
Raw gain	1.51	1.48	1.95	2.80	13.94	0.003
Residual gain	0.66	1.24	1.48	2.10	25.79	0.000
Standardized gain	2.00	1.19	1.20	1.27	8.10	0.044
Corrected residual gain	--	--	--	--	--	---

[a] Probabilities reported are exact, rounded off to the nearest thousandth.

[b] In the absence of a reliability estimates, variance of corrected residual scores can not be computed.

of attitudes and behavior. To summarize the data on personal demographics:

1. Black workers displayed more variance of raw gain scores for five measures of worker attitudes and behavior than white workers. This is a result of the greater variance found for black workers at Phase I, in 1972-73. National survey data indicated similar effects in 1969 that had diminished by 1973. Evidently some historical change had occurred, such that the difference in variance among blacks and whites was decreasing from 1969 (the first national survey) through 1972-73 (second national survey and Phase I here), to late 1974 (Phase II in the present study). If such within-phase differences do not reappear, no difference in variance of gain scores of black and white attitudes should be expected. Black workers were also found to display more variance (for two gain score types) for QoE/Comfort.

2. Women displayed more gain score variance than men for three measures of working conditions. Men displayed more variance over time for JS/Coworkers.

3. Tenure was found to be a source of heteroscedasticity for 10 of the 17 work-related variables. Typically, the middle range of tenure studied (3-5 years) had the most variance over time and had the lowest autocorrelations of the variables studied. It is also true, however, that the middle range of tenure had more blacks, more women and more low income workers than the other two tenure groups. This confounding may help explain these results, as women, blacks, and lower income workers typically displayed more variance over time.

4. Age was found to be a source of heteroscedasticity for 12 measures. Raw gain and residual gain score variances were more affected

than standardized gain or corrected residual gain score variance by age level. The younger workers (22-29) tended to have the most variance of changes and the middle age group (30-44) tended to have the least variance of changes.

5. Educational level was a source of heteroscedasticity of change score variances for all but one of the six measures of working conditions, and for six of the 11 attitudinal and behavioral measures. Raw gain and residual gain score variance were affected by educational level more than standardized gain and corrected residual gain score variance.

For the measures of working conditions, the highest educational level studied (some college but no degree) tended to have the least variance over time, and the lowest educational level (1-8 years) tended to have the most variance over time. Among worker attitudes and behavior, workers with some high school tended to have the least variance over time.

5 *Summary of results on the identification of situations with varying stability of scores.* Gross job change has been found to increase the variance over time for certain measures of working conditions, as well as for involvement and JS/Coworkers. Most of these increases seem due to greater variance of reported working conditions and involvement among the transferred workers while in their new jobs. The greater variance of reported working conditions among the transferred workers in their new jobs may, in turn, be attributed not so much to their being in particularly *varied* jobs as to their being in *new* jobs. As such, they have not yet conformed to the norms of their coworkers in regard to how their jobs are described (e.g., Job A is "known" to be repetitive, Job B is "known" to have adequate resources, etc.). One of the most direct measures of the presence of a group norm is the

reduced variance of the behavior in question among group members (e.g., Coch & French, 1948), and conversely, the absence of a group norm may be observed in the increased variance of behavior among group members. One way to test this explanation of increased variance of ratings of working conditions as resulting from the "newness" of the jobs rather than the objectively varied character of the jobs is available in another group of workers in new jobs: the subsample of replacement workers (i.e., gross person change). Since these workers replace another group (the turnover workers) in the same jobs, the objective variation between jobs is controlled. The increased variance in ratings of all six measures of working conditions is indeed replicated in the group of replacement workers (see Table 9). and, as was true of the transfer workers, such differences in variances are not found for the measures of attitudes and behaviors (i.e., differences were observed in only four of the 11 measures of attitudes and behaviors [see Table 9]). Seashore's (1954) work suggests that an intervening link between job tenure and variance of member judgments is in group cohesiveness. He demonstrates that job tenure is related to increased cohesiveness, which is in turn related to reduced variance of judgments, though he does not study the direct impact of tenure on judgment variance. Group cohesiveness is, of course, related to pressures to uniformity within groups (e.g., Back, 1951; Gerard, 1954).

Gross person change increases variance over time for most of the variables and most of their change scores. The turnover and replacement of a worker in a job may be expected to increase error variance scores, not only because the replacement workers are more varied in their reactions to the work place than their precedessors and the continuing workers, but also because their reactions to the work place are not predictable on the basis of the reactions of their predecessors, i.e., the interphase correlations of measures were typically low. Under gross job change (above), though variance was increased, predictability of perceptions and attitudes from the workers' old jobs to their new jobs was only slightly, if at all, lower than the predictability for the continuing workers.

Low income jobs usually displayed more variance over time than higher income jobs, and clerical jobs tended to display less variance over time than the other occupational groups studied. Black workers and female workers displayed more variance over time for a number of the measures investigated. Workers with three to five years of tenure tended to display more variance over time than those with more or with less tenure, for 10 work-related variables. These workers, however, earned less and had a larger proportion of women and blacks than the other categories of tenure. Such confounding makes interpretation of this result concerning tenure more difficult.

The younger workers studied tended to have the most variance over time and those aged 30-44 tended to have the least among the 12 variables that displayed age-related heteroscedasticity of gain scores.

Relatively well educated workers were more likely to be the most stable of the educational groups for the measures of working conditions, and those least educated displayed the most variance of changes over time. For the six attitudinal and behavioral measures displaying heteroscedasticity of change scores, those with some high school tended to have the least variance of change scores.

These sources of heteroscedasticity are not independent of each other. Race and sex are both confounded with income; age is confounded with education; and there are strong occupational group differences in education, race and income. Such confounding is not unique to this sample, but exists in many work places. Thus, a researcher dealing with a mostly female sample is also dealing with a low income sample. The information that has been presented here for different demographic variables is thus somewhat redundant. Whereas women in this study typically displayed more variance of gain scores than men, clericals (who are mostly women) tended to show less variance of change scores than did other occupational groups. Enough women in this sample were employed in occupational groups other than clericals for the effects of sex and occupational group to be distinguished.

As was true of the results reported earlier on identifying suitable covariates, unique and well specified relations have not been established. Rather, the evidence of heteroscedasticity deriving from its confounded sources has been presented. This evidence is likely to be useful in those many other situations where the demographic variables are as confounded as they are here.

4. The Stability of Structure of Relations Over Time

As regards the topic of instability of measures, it is worthwhile to investigate not only the instability of scores over time, important as it is for researchers using particular experimental or quasi-experimental designs, but also whether the interrelationships among demographic variables, working conditions, and work-related attitudes and behaviors are relatively stable across phases for those workers who retain their original jobs. In the simplest sense, the structure of these interrelationships may be said to have been preserved if the rank ordering of the strengths of the relationships between variables is retained. That is, if variable "a" is correlated more highly with variable "b" than with variable "c" at time one, this ordering of the strengths of relationships should persist at time two. An intercorrelation matrix simultaneously presents all the linear bivariate relations between the variables. The orderings of all the bivariate interrelationships may be spatially represented in a Smallest Space Analysis (SSA) map. The 23 variables represented in the SSA's in this analysis include the demographic variables of race, sex, education, tenure, income and age, and all 17 measures of working conditions, work-related attitudes and behaviors that have been used throughout this paper. The involvement measure used in the Phase II matrix is the three-item measure.

The intercorrelations of the 23 variables within each phase were obtained for the subsample of the 163 cases in which the same person remained in the same job in both phases. Since accurate SSA mappings were desired, the SSA's were prepared in six dimensions (coefficients of alienation = .056 and .071 in Phases I and II, respectively).

Schonemann and Carroll (1970) have introduced a technique that allows two such mappings to be compared: the vectors of one of the spaces are rotated to fit the vectors of the other space, thus preserving the ordering of the strength of the intercorrelations in the SSAs. This "best fit technique" was applied to the Phase I and Phase II SSAs and, after rotation, the relevant goodness of fit measures (normalized symmetric error, see Schonemman and Carroll [1970]) proved to be .091. This level of goodness of fit indicates that the two matrices are somewhat dissimilar since the normalized symmetric error (.091) is over twice the level (.042) referred to as "fairly good" by Schonemman and Carrol (1970, pp. 252-253). There would, thus, seem to be some difference between the structure of interrelations for the two phases. Tables 15 and 16 present the coordinates of the two variables in the two six-dimensional spaces: the rotated best-fit Phase I matrix (Table 15) and the target Phase II matrix (Table 16). Table 17 presents the residual matrix of distances between the two phases of the 23 variables' positions on each of the six dimensions, as well as overall squared distance measures between the variables' positions in the two phases.

The goodness of fit for the individual variables in the space was not uniform. Some variables were in almost the exact same place in the two six-dimensional spaces (e.g., JS/Challenge, QoE/Financial rewards, and QoE/Total) whereas a particularly large residual was noted for the absences variable. The squared distance of the position of the absences variable in the rotated, best fitting matrix from its target position (in the non-rotated matrix) was nearly four times as great as for the next most discrepant variable.

The absences variable was consequently removed from the matrix and new SSA mappings in six dimensions were prepared for the remaining 22 variables (coefficients of alienation were .051 and .070 for Phase I and II, respectively). When these two spaces were compared, the normalized symmetric error was reduced to .069, thus nearly halving the difference between the goodness of fit found previously of .091, and the "fairly good" level of .042 suggested by Schonemann and Carroll. Inspection of

Table 15

Coordinates in the 6-dimensional Space of Phase I Variables, Rotated to Best-fit with the Phase II Space

Variable	Dimension						Squared distance fr. origin
	1	2	3	4	5	6	
Ambiguity	0.554	1.004	-0.015	-0.374	0.005	-0.391	1.608
Involvement	0.316	-0.166	0.317	-0.747	-0.868	-0.502	1.789
Effort expended	0.092	-0.913	0.241	-0.807	-0.604	0.074	1.923
Intent to turnover	0.639	-0.232	-0.810	-0.777	-0.005	-0.251	1.786
JS/Coworkers	-0.113	0.441	-0.705	-0.526	-0.019	-0.507	1.238
JS/Challenge	0.033	0.385	-0.132	-0.516	-0.614	-0.566	1.131
JS/Comfort	0.068	0.633	-0.244	-0.137	-0.662	-0.589	1.269
JS/Resource adequacy	0.012	0.649	-0.161	-0.106	-0.353	-0.837	1.284
JS/Financial rewards	0.266	0.165	-0.569	-0.320	-0.686	-0.916	1.834
QoE/Total	-0.235	0.046	-0.189	-0.456	-0.583	-0.729	1.173
QoE/Resource adequacy	0.097	0.561	-0.175	-0.301	-0.244	-1.051	1.610
QoE/Financial rewards	-0.725	0.129	-0.295	-0.599	-0.973	-0.635	2.338
QoE/Challenge	-0.184	-0.260	-0.098	-0.819	-0.502	-0.613	1.409
QoE/Comfort	-0.192	0.166	-0.532	0.425	-0.783	-0.677	1.599
Global job satisfaction	0.342	0.396	-0.789	-0.499	-0.704	-0.218	1.688
Supervisory ratings	0.500	-0.232	-0.011	-0.131	-0.310	-0.871	1.177
Absences	0.018	-0.283	0.191	0.641	-0.058	-0.092	0.539
Age	0.994	-0.154	-0.301	0.043	-0.825	-0.345	1.904
Sex	0.404	-0.989	-0.182	-0.446	0.136	-0.844	2.104
Education	-0.946	-0.211	0.247	-0.176	-0.225	-0.033	1.083
Race	-0.043	-0.787	-0.720	-0.260	-0.039	-0.009	1.207
Income	-0.284	-0.821	-0.461	-0.379	-0.506	-0.574	1.698
Tenure	0.793	-0.732	-0.213	-0.096	-0.834	-0.288	1.997

Table 16

Coordinates in 6-dimensional Space of the Phase II Variables (Target Matrix)

Variable	Dimension 1	2	3	4	5	6	Squared distance fr. origin
Ambiguity	0.153	0.560	-0.844	-0.769	-0.309	-0.843	2.447
Involvement	0.593	-0.426	0.231	-0.573	-0.782	-0.627	1.519
Effort expended	-0.112	-0.533	-0.142	-1.000	-0.841	-0.239	2.081
Intent to turnover	0.527	-0.133	-0.854	-0.112	-0.397	-0.163	1.221
JS/Coworkers	0.214	0.288	-0.246	-0.834	-0.043	-0.675	1.342
JS/Challenge	0.179	0.145	-0.106	-0.430	-0.533	-0.512	0.795
JS/Comfort	0.315	0.176	-0.328	-0.178	-0.342	-0.496	0.632
JS/Resource adequacy	0.010	0.352	-0.312	-0.319	-0.445	-0.670	0.970
JS/Financial rewards	0.293	0.016	-0.160	-0.161	-0.845	-0.823	1.529
QoE/Total	-0.106	0.052	-0.207	-0.315	-0.670	-0.660	1.041
QoE/Resource adequacy	-0.022	0.459	-0.419	-0.379	-0.480	-0.950	1.663
QoE/Financial rewards	-0.681	0.175	-0.225	-0.597	-0.976	-0.590	2.202
QoE/Challenge	-0.078	-0.162	0.057	-0.459	-0.600	-0.497	0.853
QoE/Comfort	-0.029	-0.021	-0.509	0.251	-0.816	-0.579	1.324
Global job satisfaction	0.375	0.121	-0.274	-0.215	-0.731	-0.381	0.956
Supervisory rating	0.361	-0.414	-0.173	-0.235	-0.544	-0.929	1.546
Absences	-0.484	-0.531	-1.000	-0.524	-1.000	-0.901	3.603
Age	1.000	-0.274	-0.462	-0.201	-0.930	-0.722	2.715
Sex	0.448	-1.000	-0.264	-0.588	-0.199	-0.878	2.427
Education	-1.000	-0.552	-0.470	-0.285	-0.445	-0.585	2.147
Race	0.104	-0.891	-0.731	-0.327	-0.188	-0.366	1.615
Income	-0.162	-0.650	-0.058	-0.037	-0.318	-1.000	1.555
Tenure	0.845	-0.322	-0.797	-0.797	-0.937	-0.799	3.604

Table 17

Residual Matrix of Distances between the Variables' Positions in the Two 6-dimensional Spaces

Variable	Dimension						Squared distance between positions
	1	2	3	4	5	6	
Ambiguity	-0.401	-0.444	-0.829	-0.395	-0.314	-0.452	1.505
Involvement	0.277	-0.260	-0.086	0.174	0.086	-0.125	0.205
Effort expended	-0.204	0.380	-0.383	-0.193	-0.237	-0.313	0.524
Intent to turnover	-0.112	0.099	-0.044	0.665	-0.392	0.088	0.628
JS/Coworkers	0.327	-0.153	0.459	-0.308	-0.024	-0.168	0.465
JS/Challenge	0.146	-0.240	0.026	0.086	0.081	0.054	0.096
JS/Comfort	0.247	-0.457	-0.084	-0.041	0.320	0.093	0.389
JS/Resource adequacy	-0.002	-0.297	-0.151	-0.213	-0.092	0.167	0.193
JS/Financial rewards	0.027	-0.149	0.409	0.159	-0.159	0.093	0.250
QoE/Total	0.129	0.006	-0.018	0.141	-0.087	0.069	0.050
QoE/Resource adequacy	-0.119	-0.102	-0.244	-0.078	-0.236	0.101	0.156
QoE/Financial rewards	0.044	0.046	0.070	0.002	-0.003	0.045	0.011
QoE/Challenge	0.106	0.098	0.155	0.360	-0.098	0.116	0.197
QoE/Comfort	0.163	-0.187	0.023	-0.174	-0.033	0.098	0.103
Global job satisfaction	0.033	-0.275	0.515	0.284	-0.027	-0.163	0.450
Supervisory rating	-0.139	-0.182	-0.162	-0.104	-0.234	-0.058	0.147
Absences	-0.502	-0.248	-1.191	-1.165	-0.942	-0.809	4.630
Age	0.006	-0.120	-0.161	-0.244	-0.105	-0.377	0.253
Sex	0.044	-0.011	-0.082	-0.142	-0.335	-0.034	0.143
Education	-0.054	-0.341	-0.717	-0.109	-0.220	-0.552	0.999
Race	0.147	-0.104	-0.011	-0.067	-0.149	-0.357	0.187
Income	0.122	0.171	0.403	0.342	0.188	-0.426	0.540
Tenure	0.052	0.410	-0.584	-0.701	-0.103	-0.511	1.276

the squared distances between the 22 variables' positions in the two spaces revealed no further notably discrepant variables.

The actual correlations between "absences" and the other variables were inspected for both phases to identify which of its relationships had changed. These differences are considered in order of size. The largest difference concerns the correlation between absences and supervisory rating which changed from 0.34 and Phase I to 0.05 at Phase II. The correlations with supervisory ratings were computed, however, on a sample smaller than that used for the rest of the matrix, since supervisory ratings were available for only 36 of the 163 workers in the same job in both phases. Owing to the small sample size, the difference in correlation of absences with supervisory rating between Phase I (0.34) and Phase II (0.05) is not significant (Fisher's r to z, z diff = 1.23, $p > 0.20$). The next largest difference in absences' correlations over time with another variable concerned QoE/Financial rewards ($r=-0.06$ at Phase I, $r=0.19$ at Phase II, Fisher's r to z, z diff = 2.20, $p < 0.04$). Yet, one such significant difference among 22 pairs of correlations could be expected by chance.

The third largest difference in absences correlation over time with another variable was with tenure, but the difference of .16 between the intercorrelations for the two phases proved non-significant.

A later section of this paper argues that the tenure variable has different characteristics in the two phases of measurement. The changing correlation of absences with tenure may, in consequence, be meaningfully understood in terms of the changing nature of the tenure variable, even though this difference does not attain statistical significance.

Given that the tenure variable changed from Phase I to Phase II, the successful rotation of tenure in the comparison procedure to about the same position it had occupied in the previous space may have been performed at the expense of the absences variable, which was thus forced into very different positions in the two spaces. The notion that the discrepancy in absences' positions was attributable to the changing character of the tenure variable may be tested by preparing SSAs without the tenure variable but with absences included. Such SSA's were prepared (coefficients of alienation = .059 and .061 for Phases I and II, respectively) and when compared yielded a normalized symmetric error of .071. The improvement of goodness of fit is about the same improvement as had been achieved when absences had been removed (then, normalized symmetric error was .069). With both of these variables removed from the matrix, and with SSAs prepared in six dimensions (coefficients of alienation = .057 and .064 for Phases I and II, respectively), there was no further improvement in the goodness of fit measure beyond that obtained by the removal of either one alone (normalized symmetric error = .071). Such results suggest that it is the changing intercorrelation of tenure and absence that disturbs the goodness of fit rather than some aberration in the structure of absences' relationships to other variables.

The discussion of the changes in structure of relations over time has thus far concentrated on the changing positions and meaningfulness of two variables, tenure and absences. Without one or both of these variables in the matrices, however, the goodness of fit measure still did not drop to the .042 level indicated as desirable by Schonemann and Carroll. It would seem, then, that at least moderate change occurs in the structure of relations over time among the same workers in the same jobs. Yet, overall, these changes seem inconsequential and are probably a result of the sampling variance of r. The standard error of r for the sample size here is approximately 0.08, and, thus, random variations in the r's sampled from one time or another might account for the discrepancies in the rank ordering of the correlations from one time to the next.

Tenure. The difference between the correlations of absences with tenure in the two phases may be attributable to the fact that the tenure variable includes values of one year or more at Phase I, but is limited to values of three years or more at Phase II.

The limitation of the range of the tenure variable in Phase II is due to the passage of approximately 20 months, the interim between measurement of the same workers studied over time. Since the first years on the job involves a great deal of learning and uncertainty (one is then a "new" worker), the impact of job tenure need not be a linear function of the number of years on the job. The difference between being in the same job for one year and three years may, thus, be more important than the difference between three and five years, or 10 and 12 years. Hence, the tenure variable at Phase II suffers from a severe restriction of its psychologically meaningful range since the differences between one, two, and three years are simply no longer in the range of the tenure variable. The effects of this restriction of range are evident from the sizes of correlations of tenure with other variables at Phase II, as compared to Phase I, in the sample of workers in the same jobs in both phases. In Phase I, seven of the

17 correlations of tenure with the non-demographic variables were significant at the .05 level, compared to only three in Phase II, a finding consistent with the notion of restriction of range of tenure in Phase II.

Summary of results on stability of structure over time.

1. Over the period studied and among these U.S. workers:
    a. maturational changes (from ages 19 and 60 at Phase I to ages 21 and 63 at Phase II) do not seem to have had much impact on the structure of interrelations;
    b. historical changes from 1972-73 to late 1974 do not seem to have had much impact on the structure of interrelationships. Quinn, Mangione and Mandilovitch (1973) similarly report little change in the relationships between various demographic variables and reported working conditions among U.S. workers between 1969 and 1973.

2. The scale of number of years on the job (tenure) does not seem to act as an interval scale with respect to such relevant psychological constructs as task familiarity and role adaptation. At least one major distinction, that of new workers versus old timers, is not adequately represented by the chronological number of years of tenure.

General Discussion

The results from this study indicate that an experimenter who uses a non-equivalent control group design to study working conditions is in a relatively fortunate position, since under "control group" conditions there seems to be relatively little spontaneous variation that is not attributable to measurement error. The variance of changes, the error term in this experimental design, may even be less than would be found in an ordinary between-subject design, at least for populations of workers and jobs studied under comparable time lags to those in the present study.

Limits to the generalizability of these findings on the stability of working conditions over longer lags and in other cultures are suggested by two studies. Farris (1969) records considerably less stability than found here on two measures of working conditions over a lag of six years. With such longer lags, the error term of an experimental study would not be less than in an ordinary, between-subjects design. Butterfield and Farris (1974) report autocorrelations of working conditions' measures (The Likert Organizational Profile) over an approximately 12 month lag in Brazilian development banks. Their data indicate less stability of working conditions than was found in the present sample of American workers located in a hospital and two manufacturing institutions.

The researcher studying worker attitudes and behaviors would seem to be less fortunate, since these variables display considerable spontaneous variance over time above and beyond variance resulting from errors of measurement. The error variance terms for these variables are generally larger than those in a between-subjects design, and sample sizes would need to be increased accordingly. Over longer lags, Farris (1969) finds lower autocorrelations of supervisory ratings of engineers than were found for supervisory ratings in the present study. Sample sizes would need to be even greater with such lags, unless other steps were taken to reduce the error variance term.

Of the easily available covariates, namely the demographic variables, age stands out on account of its remarkably consistent correlations with a wide range of measures of working conditions, worker attitudes and behaviors. Age also has strong effects on the variance of the work-related measures. All the demographic variables studied (income, occupational group, race, sex, tenure, age and education) display heteroscedasticity of gain scores of at least a few work-related variables, although the heteroscedasticity found for race may no longer be present. With careful control of possibly confounding demographic variables, the effects for tenure also might not generalize to other investigations.

Worker age would seem, however, to be not only an empirically useful variable, but also a variable deserving of substantive attention. The processes associated with aging as well as those associated with increased tenure (a variable that appears to display changing characteristics over its scale values) might in fact best be studied in an analysis over time, the very analysis necessitated by non-random assignment of research subjects and the consequent non-equivalent control group. The study of variables over time may, therefore, become not just a solution to a specific methodological problem, but also a source of new knowledge difficult to acquire in any other way.

A researcher may feel that large amounts of error variance might be added by the replacement of experimental subjects between pre- and post-test, or by random changes in the situations of the subjects. In the case of work-related variables, this implies a concern with turnover and replacement and with subtle or not so subtle job changes that may occur, though unrelated to the experimental manipulation.

Study of the variances over time under conditions of turnover and replacement, job transfer, and no change reveals that this concern is entirely appropriate for worker turnover. For many of the change scores, most variables display increased variance over time when the job occupants had changed from Phase I to Phase II. Much weaker effects were found for job change (by worker transfer), and for most variables (particularly the attitudinal and behavioral variables) it would seem that the researchers' concerns about job changes adding error variance are misplaced. Of course, the job changes in the sites studied here may have been unusually small, but there is no obvious reason to believe that such is the case.

The impact of job change versus person change is not only an empirical concern for a researcher worried about the error variance in his or her study; it also involves a controversy in social psychology about the relative importance for social behavior of the person versus the situation (e.g., Alker, 1972; Mischel, 1968). For the present study, this issue translates into a question concerning the relative impact on worker attitudes and worker behaviors of the characteristics of a particular worker versus the characteristics of his or her job. Some of the results already discussed bear directly on this issue.

Specifically, the correlations over time of a variable with itself for the gross job change group (i.e., when measures from the workers' previous job are correlated with measures in his or her new job) and for the gross person change group (when responses from the job occupant at Phase I are correlated with the responses of the new job occupant at Phase II) reappear in Table 18. If the job situation has the greater impact, the correlations over time of attitudes and behavior should decline more when the job has changed than when the person has changed. If the particular person in a job has the greater impact, the correlations over time of attitudes and

Table 18

Correlations Over Time of Work-related Variables with Themselves in Job Change, Person Change, and No Change Groups

	Job change	Person change	No change
QoE/Resource adequacy	.14	.09	.47
QoE/Financial rewards	.58	-.06	.61
QoE/Challenge	.63	.09	.72
QoE/Comfort	.51	.34	.46
QoE/Total	.65	.31	.66
Ambiguity	.42	-.24	.41
Global job satisfaction	.51	.22	.51
JS/Coworkers	.47	.17	.16
JS/Challenge	.63	.21	.57
JS/Comfort	.57	.48	.45
JS/Resource adequacy	.42	.36	.51
JS/Financial rewards	.56	-.02	.54
Effort expended	.53	.05	.49
Supervisory rating	-.08	-.11	.59
Absences	.32	.14	.11
Turnover	.44	.36	.35
Involvement	.37	.30	.45

behavior should be lower when the person has changed. Some researchers (e.g., Argyle & Little, 1972; Endler, 1975) contend that it is neither the person nor the situation *per se* that has the greater impact on scores; rather it is the interaction of the person and the situation. In the present instance, a "strong" interactionist position would be that every combination of a person and a particular job is unique, so that only by chance would there be similar attitudes or behaviors when either the job occupant (the person) or the job (the situation) was changed. The strong interactionist position would predict near zero correlations under *both* person change and job change conditions. The present data, unfortunately, do not also permit testing of another, weaker interactionist position according to which there is an interaction *and* a person or a situation effect.

In fact, the correlations are not typically reduced to zero under either person change or job change conditions, hence, the strong interaction position is not supported. Certain versions of the strong interactionist position (Bem & Allen, 1974; Mischel, 1973) contend, however, that part of the "uniqueness" of the person-situation combination is due to the *perceived* situation, as perceptually encoded by the particular individual. Under job change the correlations of the perceptions of working conditions are, in the new and old jobs, very similar to the correlations under no change (i.e., high correlations in both groups). Accordingly, the job change does not seem to be perceived as much of a change, so the fact that the autocorrelations of attitudes and behaviors were not reduced to zero under job change is not damaging to these varieties of the interactionist position. The remaining test of the interactionist position (that the correlations should reduce to zero under person change) remains, unfortunately, virtually indistinguishable from an extreme version of the "person" approach.

Returning then to the issue of the relative importance of the person versus the situation, we observe that the correlations are consistently lower under conditions of person change, hence the question: "Does this really result from the greater importance of the person, or can some alternative explanation be offered for these results?".

These results are, unfortunately, somewhat ambiguous. The job changes (situation effect) are no more extreme than one would ordinarily find in work organizations, e.g., a clerical becomes an administrator, a laborer becomes a custodial worker, or an in-patient doctor becomes a night shift doctor in the emergency room. Such job changes as from a doctor to a nurse, a laborer to a food technician, or a clerical to a machine maintenance worker are not likely to be observed. In other words, the job changes are typically not large. The fact that the correlations of worker attitudes and behaviors between their old and new jobs are not lower than the correlations found under "no change" conditions could, then, be attributed to the slight changes involved in job change, rather than to the unimportance of situational effects, *per se*. Further discussion is accordingly limited to comparisons of the "person change" and "no change" columns of Table 18.

Person change has a considerable impact on worker attitudes and behaviors since the correlations under person change are strikingly reduced from the values found under "no change" conditions (the median of the 11 attitude and behavior measures' autocorrelations under person change is 0.21, whereas under "no change" conditions the median autocorrelation is 0.49). Because the autocorrelations under person change are not reduced to zero (in fact, despite the small sample size of the person change group, three of the 11 attitude and behavior autocorrelations are still significant at the 0.05 level two-tailed), it would seem that the different workers' situations in their different jobs also had some impact on the workers' attitudes and behaviors in addition to the substantial effects on the "person" variable. In regard to the "perceived situation" approach to the strong interactionist position (Bem & Allen, 1974; Mischel, 1973), the fact that the correlations were not reduced to zero would be explicable if there were consistency in the perceived situations of the different job occupants.

When the autocorrelations under "person change" and "no change" conditions are compared for the six measures of working conditions, the effects found for person change are much stronger and the situation effects much weaker than was found for the worker attitudes and behaviors. The median autocorrelation of the six measures of working conditions under person change is 0.09, whereas under "no change" conditions, the median autocorrelations is 0.54. One would have expected that the situational differences would have had greater impact on perceptions of working conditions than on attitudes and behavior, yet it is found that changing the occupant of the very same job is sufficient to reduce the autocorrelations of the ratings of working conditions to practically zero! This indicates that the consistency of attitudes and behaviors under person change was not attributable to the consistency in the preceived working conditions of the different job occupants, as the "perceived situation" approach to the strong interactionist would have it. Evidently there are simply powerful "person" effects on attitudes, behavior, and perceived working conditions. The person effects are stronger (and the objective situational effects weaker) for measures of working conditions than for measures of attitudes and behavior.

One can only conclude that these measures of working conditions are not measuring to any great extent objective characteristics of the job that exist independently of the particular occupant of the job who reports these working conditions. More generally, Golding (1977) indicates that only 25% of a typical environment scale's variance (for any environment, not just work situations) reflects between environment differences. Similar findings may be noted with another set of worker rated measures of working conditions - Hackman and Oldham's (1975) job description scales. Using Hays' (1963, p. 382) formula for a rough estimate of the percentage of variance in ratings of jobs explained by actual job differences on the data presented by Hackman and Oldham (1975, Table 3), we find that a median value of only 16% of the variance of the seven subscales is explained by job differences. In our data, we may approach the correlations between ratings of working conditions by different occupants of the same jobs as measures of interrater reliability. The median interrater correlation for the measures of working conditions of 0.09 indicates that a median value of 9% of the variance of rated working conditions is explained by job differences, a finding not too different from the median of 16% found for Hackman and Oldham's seven subscales. This difference may reflect the passage of nearly two years between ratings in our data.

The results suggest that the measures of working conditions obtained reflect at least two processes: the development of a group norm characterizing working conditions which influence the variance of ratings (noted earlier in the summary of Section 3 of the Results and Discussion); and some individual difference variable which is expressed in the ordering of the ratings of working conditions offered by different people.

The nature of this individual difference variable is not clear. Different people may rate working conditions differently because of an "implicit theory of working conditions", much like an "implicit theory of leadership" (Eden & Leviatan, 1975). People may carry around their "theory" in their heads and, accordingly, fill out questionnaires in a way that is relatively independent of the "objective" task characteristics. Another way to understand this individual difference variable is to look at the individual's influence on his or her job. Thus, for example, a certain worker may always demand freedom from supervision on the job, so that no matter what the history of supervision on the job, this supervision very quickly disappears. This notion is consistent with the "transactional" interaction approach that recognizes reciprocal causation of the person and the situation (Bowers, 1973; Overton & Reese, 1973; Perrin, 1968). Jenkins and Nadler (Chapter 7) note that Hackman (1969) hypothesizes the existence of a process of job redefinition on the part of the worker. Redefinition may be "objective", as in the example given above of actual changes in supervisory behavior, or subjective in that the worker distorts his or her perceptions of the job to be consistent with his or her wants, needs, goals, and values. Jenkins and Nadler also note that Alderfer (1972) proposes a similar process to Hackman's (1969) job redefinition notion. Whether either of these two interpretations (implicit theory or job redefinition) is the source of the individual difference variable here is not clear, but its effects are quite compelling in view of the low correlations found between ratings of working conditions of the very same jobs by different occupants.

It was suggested above that the differences between jobs observed under the "job change" conditions are typically not very great, so that the reduction of correlations from the values under conditions of "no change" would not be particularly large, and would not be a fair test of the power of situational effects upon worker attitudes and behavior. Some reduction of the attitude and behavior autocorrelations would, however, still be expected under job change as compared to no change conditions. It is interesting to note, then, that the autocorrelations under job change of attitudes and behaviors were reduced from their values in "no change" conditions for only three of the 11 attitude and behavior variables. It is difficult to see why changing jobs should increase the correlations of attitudes and behaviors between phases as compared to not changing the jobs between phases. Perhaps instead of viewing the difference in correlations as resulting from changing the condition, it would be more fruitful to ask why not changing the job would reduce correlations over time.

For those in the no change group, staying in the same job doubtless increased their familiarity with their jobs. For some of these workers the increased familiarity may have increased their jobs' attractiveness: Zajonc (1968) reviews several studies which demonstrate that increasing familiarity of an attitudinal object increases the object's attractiveness. For other workers the increased familiarity likely induced boredom and so reduced the attractiveness of their jobs. Thus, the very lack of job change may modify attitudes and behaviors in different directions for different people, thereby creating a form of interaction between person and familiarity effects. Katerberg, Smith, and Hoy (1977) note this kind of interaction over a time span of only six weeks. Such an interaction would reduce the linear correlations over time for the no change group as a whole. A similar argument was used earlier (Section 2C of Results and Discussion) in regard to the reliability of change, where it was suggested that the very absence of reliable change in working conditions might be related to the reliable changes found for worker attitudes and behaviors. We have, then, in the no change group (as opposed to the job-change group) a discrepancy between the extent of perceived similarity of the jobs in the two phases and the extent of similarity of reactions (the attitudes and behaviors) to the jobs in the two phases. Magnusson and Ekehammer (1975) have also observed such discrepancies for certain classes of situations.

Future panel studies will, hopefully, further clarify the interrelationships of worker attitudes, behaviors, working conditions, and individual difference factors, as their dialectical interplay over time (Riegel, 1976) is revealed.

APPENDIX 1

The variance of the raw gain score ($var[y_i - x_i]$, where x_i is the score at time one and y_i the score at time two for the i'th case), by using the algebra of expectations may be shown to equal:

VAR(Y) = VAR(X) - 2COVAR(X,Y)

Substituting the appropriate sample estimates, we find that this equals:

(1) $s_y^2 + s_x^2 - 2r_{xy}s_xs_y$.

Substituting into this formula the variance for standardized variables (where $s_x^2 = s_y^2 = 1$) to find the variance of standardized gain, since 'r' is invariant under linear transformation, we find that formula 1 reduces to:

(2) $2(1-r_{xy})$.

The variance of residual scores is, of course:

(3) $(1-r_{xy}^2)s_y^2$.

The variance of corrected residual scores may be found by substituting the appropriate values into formula 1. We wish to find

$$\mathrm{VAR}\left(y_i - \left[\frac{b_{y.x}}{r_x} \right] x_i \right)$$

where $b_{y.x}$ is the raw linear regression coefficient in predicting y_i from x_i, and where r_x is the reliability of the Phase I measure. Substituting the symbol y^1 for $\frac{b_{y.x}}{r_x} x_i$, we find that:

$$\mathrm{VAR}\left(y_i - \frac{b_{y.x}}{r_x} x_i \right) = s_y^2 + \frac{b_{y.x}^2 s_x^2}{r_x^2} - \frac{2r_{yy^1} s_y s_x b_{yx}}{r_x} .$$

Since y^1 is a linear transformation of x, then:

$$r_{yy^1} = r_{xy}.$$

It is also true that:

$$b_{y.x} = \frac{r_{xy}s_y}{s_x}$$

Substituting terms and canceling, we find that the variance of corrected residual scores is:

$$s_y^2 + \frac{r_{xy}^2 s_y^2}{r_x^2} - \frac{2r_{xy}^2 s_y^2}{r_x} \; .$$

APPENDIX 2

The letters x and y will be used to represent the initial score and the final score respectively, r_x and r_y the internal consistency coefficients, and r_{xy} the stability coefficient across the interval of observation. s_x and s_y are the standard deviations of the initial scores and final scores respectively.

$$\text{Reliability of Raw Gain} = \frac{r_y s_y^2 - 2r_{xy} s_x s_y + r_x s_x^2}{s_y^2 - 2r_{xy} s_x s_y + s_x^2}$$

$$\text{Reliability of Standardized Gain} = \frac{r_y + r_x - 2r_{xy}}{2(1-r_{xy})}$$

$$\text{Reliability of Residual Gain} = \frac{r_y - 2r_{xy}^2 + r_{xy}^2 r_x}{1 - r_{xy}^2}$$

$$\text{Reliability of Corrected Residual Gain} = \frac{r_x(r_x r_y - r_{xy}^2)}{r_x^2 - 2r_{xy}^2 r_x + r_{xy}^2}$$

These formulae are taken from Davidson (1972), Lord (1963) and Tucker et al. (1966).

References

Alderfer, C. P. *Existence, relatedness, and growth: Human needs in organizational settings*. New York: The Free Press, 1972

Alker, A. A. Is personality situationally specific or intrapsychically consistent? *Journal of Personality*, 1972, *40*, 1-16

Argyle, M., & Little, B. Do personality traits apply to social behavior? *Journal of the Theory of Social Behavior*, 1972, *2*, 1-33.

Back, K. W. Influence through social communication. *Journal of Abnormal and Social Psychology*, 1951, *46*, 9-23.

Barnowe, J. T., Mangione, T. W., & Quinn, R. P. *The relative importance of job facets as indicated by an empirically derived model of job satisfaction*. Ann Arbor: Survey Research Center, 1972 (mimeo).

Beehr, T. A. Role ambiguity as a role stress: Some moderating and intervening variables (Doctoral dissertation, The University of Michigan, 1974). Dissertation Abstracts International, 1974, *35*, 3631B-3632B. (University Microfilms No. 75-630, 170.)

Beehr, T. A. Perceived situational moderators of the relationship between subjective role ambiguity and role strain. *Journal of Applied Psychology*, 1976, *61*, 35-40.

Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 1974, *81*, 506-520.

Bowers, K. S. Situationism in psychology: An analysis and a critique. *Psychological Bulletin*, 1973, *80*, 307-336.

Butterfield, D. A., & Farris, G. F. The Likert Organizational Profile: Methodological analysis and test of system 4 theory in Brazil. *Journal of Applied Psychology*, 1974, *59*, 15-23.

Campbell, D. T., & Stanley , J. D. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally & Co., 1963.

Cobb, W., Jr. *The relationship between quality of employment and job satisfaction among black and white workers*. Ann Arbor: Survey Research Center, March 1973 (mimeo).

Coch, L, & French, J. R. P., Jr. Overcoming resistance to change. *Human Relations*, 1948, *1*, 512-532.

Cronbach, L. J., & Furby, L. How should we measure 'change'-- or should we? *Psychological Bulletin*, 1970, *74*, 68-80.

Davidson, T. N. *Youth in transition*. *Volume IV*. *Evolution of a strategy for longitudinal analysis of survey panel data*. Ann Arbor: Institute for Social Research, 1972.

Duncan, O. D. Path analysis: Sociological examples. American Journal of Sociology, 1966, 72, 1-16.

Eden, D., & Leviatan, U. Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. Journal of Applied Psychology, 1975, 60, 736-746.

Endler, N. S. The case for person-situation interactions. Canadian Psychological Review, 1975, 16, 12-21.

Farris, G. F. Organizational factors and individual performance: A longitudinal study. Journal of Applied Psychology, 1969, 53, 87-92.

Fleiss, J. C. Comment on Overall and Woodward's asserted paradox concerning the measurement of change. Psychological Bulletin, 1976, 83, 774-775.

Gerard, H. B. The anchorage of opinions in face-to-face groups. Human Relations, 1954, 7, 313-326.

Glickman, A., & Brown, Z. Changing schedules of work: Patterns and implications. Washington, D.C.: American Institutes for Research,

Golding, S. L. The problem of construal styles in the analysis of person-situation interactions. In D. Magnusson and N. S. Endler (Eds.), Personality at the crossroads: Current issues in interactional psychology. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Hackman, J. R. Toward understanding of the role of tasks in behavioral research. Acta Psychologica, 1969, 31, 97-128.

Hackman, J. R., &. Oldham, G. R. Development of the job diagnostic survey. Journal of Applied Psychology, 1975, 60, 159-170.

Harris, C. W. (Ed.). Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Katerberg, R., Smith, F. J., & Hoy, S. Language, time, and person effects on attitude scale translations. Journal of Applied Psychology, 1977 62, 385-391.

Kenny, D. A. A quasi-experimental approach to assessing treatment effects in the non-equivalent control group design. Psychological Bulletin, 1975, 82, 345-362.

Kessler, R. C. The use of change scores as criteria in longitudinal survey research. Quality and Quantity, 1977, 11, 43-66.

Lord, F. M. Elementary models for measuring change. In C. W. Harris (Ed.), Problems in measuring change. Madison: University of Wisconsin, 1963.

Magnusson, D. & Ekehammer, B. Perceptions of and reactions to stressful situations. Journal of Personality and Social Psychology, 1975, 31, 1147-1154.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Mischel, W. Toward a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Overall, J. E., & Woodward, J. A. Unreliability of difference scores: A paradox of measurement of change. Psychological Bulletin, 1975, 82, 85-86.

Overall, J. E., & Woodward, J. A. Reassertion of the paradoxical power of tests of significance based on unreliable difference scores. Psychological Bulletin, 1976, 83, 776-777.

Overton, W. F., & Reese, H. W. Models of development: Methodological implications. In J. R. Nesselroade & H. W. Reese (Eds.), Life span developmental psychology: Methodological issues. New York: Academic Press, 1973.

Pelz, D. C., & Andrews, F. A. Scientists in organizations. New York: Wiley, 1966.

Perrin, L. A. Performance and satisfaction as a function of individual-environment fit. Psychological Bulletin, 1968, 69, 56-68.

Quinn, R. P., Mangione, T. W., & Mandilovitch, M. Baldi de. Evaluating working conditions in America. Monthly Labor Review, 1973, 96(11), 32-41.

Quinn R. P, Seashore, S. E., Kahn, R. L., Mangione, T. W., Campbell, D. B., Staines, G. L., & McCullough, M. Survey of working conditions: Final report on univariate and bivariate tables (Document No. 2916-0001). Washington, D.C.: U.S. Government Printing Office, 1971.

Quinn, R. P., & Shepard, L. J. The 1972-73 quality of employment survey: Descriptive statistics with comparison data from the 1969-1970 survey of working conditions. Ann Arbor: Survey Research Center, 1974

Riegel, K. F. The dialectics of human development. American Psychologist, 1976, 31, 689-700.

Schonemann, P. H., & Carroll, R. M. Fitting one matrix to another under choice of a central dilation and a rigid motion. Psychometrica, 1970, 35, 245-255.

Seashore, S. E. Group cohesiveness in the industrial work group. Ann Arbor: Institute for Social Research, 1954.

Tucker, L. R., Damarin, F., & Messick, S. A base-free measure of change. Psychometrica, 1966, 31, 457-473.

Werts, C. E., & Linn, R. L A general linear model for studying growth. Psychological Bulletin, 1970, 73, 17-22, (a).

Werts, C. E., & Linn, R. L. Path analysis: Psychological examples. Psychological Bulletin, 1970, 74, 193-212, (b).

Work in America (Report of a special task force to the Secretary of Health, Education and Welfare.) Cambridge: The MIT Press, 1973.

Zajonc, R. B. The attitudinal effects of mere exposure. Monographs of Journal of Personality and Social Psychology, 1968, 9 (2, Part 2) 2-27.

Chapter 9

AGGREGATING DATA IN ORGANIZATIONAL RESEARCH

by

Nina Gupta and Robert P. Quinn

ABSTRACT

Organizational studies vary considerably in terms of (1) whether or not they have aggregated data obtained from individuals (e.g., on a work group basis), and (2) their choices of bases of aggregation. Data obtained from 651 employees of five firms were used to test the relative reliabilities of three bases for aggregating individual data: occupation, reporting to the same supervisor, and a combination of the two bases. Regardless of whether descriptions of jobs, work groups, or supervisors were being aggregated, the most reliable basis of aggregation was the combined one, followed by bases that involved reporting to the same supervisor and having the same occupation, respectively.

Chapter 9

AGGREGATING DATA IN ORGANIZATIONAL RESEARCH

The behaviors of people in organizations are sometimes investigated by using neither people nor organizations as units of analysis. For example, both turnover and absenteeism are behaviors of individual workers and could be treated analytically as such. But Lyon's 1972 review of studies of the relationship between these two behaviors indicated that about two-thirds of the studies aggregated individual turnover and absenteeism data on the basis of work groups, departments, or other administrative units and used the latter, rather than individual workers, as units of analysis. Other bodies of organizational research also vacillate in whether the unit of analysis is the individual member or some larger organizational unit. In 13 of the studies of the relationship between employees' attitudes and their job performances reviewed by Brayfield and Crockett (1955), the unit of analysis was the individual worker; in the remaining ten studies it was something approximating a work group. Vroom's 1964 review of 20 studies of the relationship between job satisfaction and job performance indicated that the individual was the unit of analysis in 15 studies and the work group in the remaining five.

If the behavior of the individual organizational member is of interest, why do some investigators not use this person as the unit of analysis? The reasons for the aggregation of data pertaining to individuals are, in order of salience in the literature, expediency, theoretical considerations, and the necessity of obtaining statistically reliable estimates.

For expediency the investigator is often faced with the annoying reality that what he or she wants to measure in a particular organization is simply not available conveniently at an individual level. For example, good measures of a person's job performance are not only difficult to obtain but are often not available in an organization's records. As a result, the investigator is compelled to rely on organizational records that provide performance data only on a work group or some other level: the scrap records of departments; the sales reports of branch offices; the number of grievances filed in different divisions. This expedient restriction to a group unit of analysis may in turn require the investigator to aggregate on the same basis other variables that need not or should not be so aggregated. If, for example, one is studying the relationship between job satisfaction and job performance in an organization and performance measures are available only for whole branch offices of the organization, one may have little choice but to aggregate employee's job satisfaction scores on a branch office basis.

For theoretical reasons the investigator is at times compelled to aggregate data. For example, Likert's (1961) conception of organizations as composed of overlapping groups with "linking pins" requires that many variables be measured at the work group level. Investigators who have adopted Likert's structural conception, instruments, methods of data collection, and analysis conventions have relied heavily upon work groups as units of analysis (Franklin, 1975a, 1975b). Occasionally they have used as units of analysis either total organizations or organizational "capstone groups" (Bowers, 1973).

Studies of organizational "climate" have differed considerably in terms of whether the unit of analysis is the individual worker, the work group, or the employing establishment. Investigators in the Likert tradition regard climate as an organizational characteristic, yet usually measure it on a work group or individual basis. Guion (1973), however, has argued that studies of organizational climate that use individual workers as units of analysis violate the whole notion that climate is ultimately "objective" and, therefore, can be estimated only at a group or organizational level. "The idea of 'organizational climate'", according to Guion, "appears to refer to an attribute, or set of attributes, of the work environment. The idea of a 'perceived organizational climate' (i.e., that wherein the individual worker is the unit of analysis) seems ambiguous; one can not be sure whether it implies an attribute of the organization or of the perceiving individual. If it refers to the organization, then measures of perceived organizational climate should be evaluated in terms of the accuracy of the perceptions. If it refers to the individual, then perceived organizational climate may simply be a different name for job satisfaction or employee attitudes" (p.120).

While one might argue whether organizational or work group climate can justifiably be assessed at the individual level, there is at least one concept that almost by definition requires estimation at the group level--group cohesiveness. It is, therefore, heartening that in Stogdill's 1972 review of studies relating productivity and group cohesiveness almost all of the studies appeared to use some kind of group as the unit of analysis.

Finally, to obtain reliable estimates of variables of interest, it is often necessary for an investigator to aggregate data. The most common example of this is the assessment of a supervisor's behavior through the reports provided by his or her subordinates. It is assumed in this instance that the supervisory behavior cannot necessarily be reported accurately by any one subordinate and that a more reliable estimate is obtainable from the averaged judgements of all those people who report directly to the supervisor.

The investigator is, therefore, sometimes confronted with the decision of whether or not to aggregate data obtained from individual organizational members into grosser units on grounds that are either expedient, theoretical, or statistical.

But what bases of aggregation are appropriate? Existing studies of organizational behavior have varied considerably in the bases used to aggregate data. Table 1 summarizes 59 studies that were published between 1970 and 1976 in six major journals specializing in studies of organizational behavior. Each of the 59 studies used units of analysis that aggregated data obtained from smaller organizational units or, most commonly, individual organizational members. Some studies in the journals reviewed were excluded from the tally in Table 1 because they failed to specify their units of analysis and provided no incidental information (e.g., n's in tables) from which the units could be inferred.

According to Table 1, the most frequently used basis of aggregation has been the primary organizational unit--usually described in the research reports reviewed as the "work group". The most common operational

Table 1

Number of Studies Using Different Bases of Aggregation as Reported in 59 Studies for 1970 through 1975[a]

Basis of aggregation	Number of studies
Total organizations	10
(Communities)	(1)
(School systems)	(2)
(Apartment complexes)	(1)
(Service agencies)	(1)
(Employing establishments)	(5)
Primary units within organizations (e.g., work teams, work groups, work pairs, all those who reported immediately to the same supervisor)	20
Secondary units within organizations (e.g., departments, divisions, bureaus, branches, offices, stations, plant sections, "units" not otherwise specified	18
"Occupational group" as idiosyncratically identified by either the investigators or by the terminologies of the organizations studied. In some instances hierarchical level in the organization was used by the investigators to identify occupational group	11

---

[a] The 59 studies were all those that used units of analysis employing aggregated data obtained from smaller organizational subunits or individual respondents and that were published from 1970 through 1975 in *Administrative Science Quarterly*, *Journal of Applied Behavioral Science*, *Journal of Applied Psychology*, *Personnel Psychology*, *Organizational Behavior and Human Performance*, or *Sociology of Work and Organizations*.

Excluded from the tally of studies using aggregated data were those wherein the units of analysis were: (1) T-groups or other groups of organizational members that did not previously exist in organizations; (2) groups of students used in experiments simulating organizational behavior; (3) a single group that was measured many times as part of an experiment.

Also excluded were studies wherein a potential basis of aggregation was used as an independent, dependent, or moderator variable. This eliminated many studies relating occupation as an independent variable to measures of workers' beliefs, attitudes, and behaviors.

definition of "work group" that could be inferred from the scant methodological descriptions provided was that it consisted of all people reporting directly to the same supervisor. Almost as prevalent have been studies wherein units of analysis have been secondary units within organizations--departments, divisions, bureaus, branches, etc. In some studies, however, sketchy reporting made it difficult to determine whether a primary or secondary organizational unit was being used to aggregate. Less common, according to Table 1, were total organizational and occupational groups as units of analysis.

A common assumption in the aggregation of data is that the basis of aggregation is meaningfully related to the commodity that is aggregated. If, for example, job content is being studied, occupation seems a meaningful basis of aggregation. If supervisors' leadership styles are being investigated, aggregating their subordinates' reports on a work group basis seems in order. Other instances are less clear. What are appropriate bases of aggregation in matrix organizations? Does not the aggregation of workers' reports of the content of their jobs on a work group basis assume that the jobs in a work group are similar, if not identical? To the extent that jobs within a work group are highly differentiated, there may be little agreement among workers' reports. The reliability of the aggregated estimates may accordingly be reduced, and their validity must suffer in the process.

The purpose of the present analysis is to assess the reliabilities of three bases of aggregating data pertaining to individual workers and their jobs: occupational categories, work groups, and a combination of the two. The variables thus aggregated were obtained both from workers' reports and independent observations of their jobs. They were chosen to represent characteristics that varied markedly in terms of whether their principal referents were job, work group, or supervision (Campbell, 1971).

Method

The data were collected from 651 employees of five organizations: a large hospital; a printing company; a research and development laboratory; and two plants that manufactured automobile accessories. Fifty-one percent of the sample were men, and 80% were white. Their average age was 35 years. Their mean income was approximately $10,000 a year, and their median education level was that of a high school graduate. Sixty percent were in blue-collar occupations. A fuller description of the sample and other methodological matters is provided in Chapter 1.

Several methods of data collection were used in the original study: (1) a 90 minute interview with each respondent; (2) two hours of observation of the respondent's job by two trained observers; (3) employers' records, and (4) supervisors' ratings of their subordinates. The present analysis uses data from only the first two of these sources. The personal interview included questions concerning the worker's job values, job descriptions, job attitudes, physical and mental health, social and political activities, personality, and personal background. The on-the-job observations attempted to describe each worker's job and job environment. The reliabilities and validities of these observations have been presented in Chapter 4.

The variables to be aggregated were chosen to represent a wide selection of those that might be useful in organizational research and that could be meaningfully aggregated on some basis. Measures of personality characteristics and attitudes, notably job satisfaction, were excluded.

Job descriptions included indicators of both quality of employment (Barnowe, Mangione, and Quinn, 1972; Quinn, 1974) and core dimensions of job content (Hackman and Lawler, 1971). The quality of employment indicators were based only upon data obtained from the interviews and included resource adequacy, financial rewards, challenge, and comfort. Hackman and Lawler's core dimensions of variety, autonomy, task identity and feedback were all measured in the interview, further differentiating feedback into that which was provided by the task itself and that provided by other sources. Variety, autonomy, and job meaningfulness were also measured by the on-the-job observations. In Chapter 4's analysis of the observational measures, these three dimensions exhibited high repeatability both at the same time and at different times.

Work group descriptions concentrated on three aspects of each respondent's report of his or her coworkers: how well they get along and stick together, which, on the group level of analysis, may represent group cohesiveness; how competent these coworkers seem; and how well they facilitate the respondent's performance on his or her job. These three measures were all obtained from the interview wherein "coworkers" were specified as follows to the respondent: "People whom you see just about every day and with whom you have to work closely in order to do your job well."

Supervisor descriptions, obtained from respondents exclusively through interviews, defined "supervisor" to each respondent as follows: "Someone who is directly over you, or someone who lets you know what you have to do." The interview questions focused upon the supervisor's supportiveness, competence, and task facilitation.

Descriptions of jobs, work groups, and supervisors were aggregated on three bases.

1. All respondents classified as having the same occupation were those who had similar three-digit occupational codes according to the Census Occupational Classifications (U.S. Bureau of the Census, 1971). Some indication of the specificity of these codes may be obtained by noting the first few occupations so coded according to the Census: aeronautical engineers, chemical engineers, entertainers, farm and home management advisors.

2. All respondents classified as having the same supervisor were those who named the same person as supervisor according to the definition already mentioned.

3. The third basis of aggregation combined the preceding two. Respondents with the same supervisor were first identified. Groups of workers so defined were then further divided into smaller groups with similar occupational codes.

A comparison between the three types of variables to be aggregated and the three bases of this aggregation suggests that the two sets may reflect somewhat similar dimensions. Job descriptions and occupational classifications indicate one such dimension, since the latter seem a logical way of aggregating the former. Likewise, having the same supervisor provides a plausible basis for aggregating data concerning work groups and supervisors. It was therefore hypothesized that for the full range of 19 work characteristics assessed the greatest within-group agreement would be obtained using the basis of aggregation that combined occupational codes and supervisor identification.

Within-group agreement--that is, the reliability of each aggregated estimate of each work characteristic--was measured by squared epsilon coefficients. Epsilon-squared is the unbiased correlation ratio and is a measure of the strength of the association made after adjusting for degrees of freedom. According to Campbell, Converse and Rodgers (1976), it can "be interpreted as the proportion of variance in the responses which can be explained by cluster; in other words, the extent of agreement among respondents within a cluster." (p.232).

Results

The squared epsilon coefficients indicating within-group agreement concerning work characteristics using the three bases of aggregation are shown in Table 2. Since each of the coefficients was limited by the internal consistency of the measure of each work characteristic, comparisons within columns of the table should be avoided. More central are the comparisons in each row among the reliability estimates obtained from the three different bases of aggregation.

Table 2

Estimates of Within-group Agreement (Epsilon-squared) on Work Characteristics Using Three Bases of Aggregation

Work characteristics[a]	Basis of aggregation[b]		
	Same occupation	Same supervisor	Same occupation and supervisor
Job descriptions			
Quality of employment			
Resource adequacy	.29	.43	.49
Financial rewards	.43	.53	.46
Challenge	.60	.63	.68
Comfort	.34	.42	.30
Overall	.59	.65	.67
Core dimensions			
Variety (interview)	.47	.51	.47
Variety (observation)	.77	.78	.84
Autonomy (interview)	.54	.51	.60
Autonomy (observation)	.78	.84	.88
Task identity (interview)	.15	.26	.31
Meaningfulness (observation)	.72	.76	.84
Task feedback	.14	.23	--[c]
External feedback	.14	.08	.33
Work group descriptions			
Cohesiveness	.25	.21	.56
Competence	.21	.34	.43
Task facilitation	.28	.38	.35
Supervisor descriptions			
Supportiveness	.38	.51	.57
Competence	.32	.45	.58
Task facilitation	.44	.60	.62

---

[a] Data were obtained from interviews unless otherwise distinguished.

[b] Ns vary between rows depending on the amount of missing data on the work characteristic. The n's averaged as follows:
same occupation: 57 groups, 567 respondents (median group size = 6);
same supervisor: 94 groups, 412 respondents (median group size = 4);
same supervisor and occupation: 68 groups, 197 respondents (median group size = 2).

[c] The correction for the number of subclasses was too large.

The best basis of aggregation was, as predicted, that which combined having the same supervisor with having the same occupation. This basis of aggregation produced the most reliable estimates of within-group agreement in 31 of the 36 comparisons with bases of aggregation that involved having only the same occupation or the same supervisor. While at times the differences between squared epsilons were large, as much as .35, they were often rather small.

The second best basis of aggregation was that based upon having a common supervisor. In 16 of 19 tests it produced a squared epsilon coefficient greater than that obtained by using occupation as a basis of aggregation.

The least reliable basis of aggregation was occupation.

There was no consistent pattern in terms of whether particular bases of aggregation provided more reliable estimates of particular variables. Workers reporting to the same supervisor consistently provided more reliable estimates of their supervisors' behaviors than did those aggregated according to their occupations. On the other hand, aggregating workers according to occupational codes failed to provide more reliable estimates of job descriptions than did aggregating on the basis of having a common supervisor. With regard to work group descriptions, neither having the same occupation nor the same supervisor seemed to yield a more reliable estimate. In short, the most reliable basis of aggregation was that which first identified a common supervisor and which then honed work groups so defined into those with similar occupations. Nearly as reliable was the identification of a common supervisor, and least reliable was occupation as embodied in Census codes.

Discussion

While, as predicted, the bases of aggregation combining workers having the same occupation and having the same supervisor provided the greatest agreement within groups, other data were contrary to prediction. Most conspicuously, occupational classifications were not the best bases for aggregating job descriptions. Although, as predicted, having the same supervisor was a good basis for aggregating workers' descriptions of their work groups and their supervisors, it also fared better than having the same occupation in aggregating job descriptions.

At least three reasons can be advanced post hoc to account for the overall superiority of aggregation on the basis of supervisor rather than occupation. First, a supervisor in an employing establishment usually has authority over a cluster of jobs that are somewhat similar and frequently identical. Therefore, having the same supervisor often means having more or less the same occupation as well. But having the same occupation in most employing establishments does not equally well insure having a common supervisor. As a result, having a common supervisor provides a better approximation of the most reliable basis of aggregation--that of combining supervisor and occupation. Second, there is considerable evidence that a supervisor's behaviors may determine certain characteristics of the jobs of his or her subordinates. Particularly within a supervisor's control are autonomy, feedback, challenge, resource adequacy, and the behaviors of one's coworkers. Aggregating on the basis of having a common supervisor therefore carries with it many implications for what one's occupation is, as well as more specific characteristics of one's job. Third, workers with a common supervisor usually work together and as a result may influence each others beliefs and attitudes. This may in turn produce greater shared standards for evaluating their jobs, and even shared evaluations, than might otherwise be the case.

The results of our tally of previous studies (Table 1) have additional implications for past and future organizational research. Studies that had aggregated data were found to be distressingly scanty in their descriptions of methods. It was not only at times unclear whether aggregation had been used, but more commonly the bases of aggregation were obscure. "Occupation" was almost never defined precisely, and what a "work group" was often failed to be specified. Such failures to clarify bases of aggregation seriously limit both a study's implications and its replicability.

The results of our analyses suggest further that the usefulness of any basis of aggregation may differ according to the organization studied. In addition to choosing the bases of aggregation to fit the purposes of the study, there is the matter of reliability. Investigators, that is, should test whether their aggregate measures are reliable enough to use as indicators of the units studied. In an organization where each work group is fairly homogenous with respect to members' jobs, for example, one would expect to fare reasonably well by using work group as the unit of analysis. But even in such an organization the reliability of aggregated estimates might possibly be low enough to require improvement by aggregation on the basis of job and supervisor. On the other hand, in an organization where jobs within each work group are heterogeneous, the combined basis of aggregation would be essential to obtain sufficiently reliable estimates. To use other bases in this case would needlessly limit the validities of the estimates and restrict the magnitudes of their correlations with other measures.

References

Barnowe, J. T., Mangione, T. W., and Quinn, R. P. The relative importance of job facets as indicated by an empirically derived model of job satisfaction. Ann Arbor: Survey Research Center, 1972 (mimeo).

Bowers, D. OD techniques and their results in 23 organizations. Journal of Applied Behavioral Science, 1973, 9, 21-43.

Brayfield, A., & Crockett, W. Employee attitudes and employee performance. Psychological Bulletin, 1955, 52, 396-424.

Campbell, A., Converse, P. E., & Rodgers, W.L. The quality of American life. New York: Russell Sage, 1976.

Campbell, D. B. Relative influence of job and supervision on shared worker attitudes. Journal of Applied Psychology, 1971, 55, 521-525.

Franklin, J. L. Down the organization: Influence processes across levels of hierarchy. Administrative Science Quarterly, 1975, 20, 153-164. (a)

Franklin, J. L. Relations among four social-psychological aspects of organizations. Administrative Science Quarterly, 1975, 20, 422-433. (b)

Glass, G. V., & Hakstian, A. R. Measures of association in comparative experiments: Their development and interpretation. American Educational Research Journal, 1969, 6, 403-414.

Guion, R. A note on organizational climate. Organizational Behavior and Human Performance, 1973, 9, 120-125.

Hackman, J. R., & Lawler, E. E. Employee reactions to job characteristics. Journal of Applied Psychology, 1971, 55, 259-286.

Likert, R. New patterns of management. New York: McGraw-Hill, 1961.

Lyons, T. Turnover and absenteeism: A review of relationships and shared correlates. Personnel Psychology, 1972, 25, 271-282.

Quinn, R. P. Strategy issues in the development of quality of employment indicators. In Proceedings of the Social Statistics Section of the American Statistical Association. Washington, D. C.: American Statistical Association, 1974.

Stogdill, R. Group productivity, drive, and cohesiveness. Organizational Behavior and Human Performance, 1972, 8, 26-43.

References (continued)

Vroom, V. Work and motivation. New York: Wiley, 1964.

U. S. Bureau of Census. 1970 census of population: Alphabetical index of industries and occupations. Washington, D. C.: U. S. Government Printing Office, 1971.

Chapter 10

JOB SATISFACTION MEASUREMENT: NORMATIVE AND INDIVIDUALISTIC MODELS

by

Nina Gupta

ABSTRACT

This paper reconsiders an old but still unresolved issue in occupational psychology. It has been proposed that the prediction of a worker's job satisfaction, from information about the various rewards and benefits provided by the job, will be improved if the worker's own sense of priorities is taken into account. The job facets viewed as most important by the worker himself should be given most weight when estimating how satisfying or dissatisfying a particular job will be. However, most previous studies of importance-weighting have been negative; importance weighting appears not to add to the predictability of job satisfaction.

A first analysis of the data shows, as expected, that the rated importance of a job facet potentially can aid in the prediction of resultant facet satisfaction. The evidence for this consists, first, of the relatively greater variance in facet satisfaction for the "high importance" workers who are known to receive much or little of the facet; that is, the high importance workers were more responsive to variations in actual facet realization than were the low importance workers. A second and confirming analysis showed that the correlations between facet realization and satisfaction are stronger for the high importance workers than for the low importance workers except for one reversal in this pattern for importance of Financial Rewards. Comparisons on mean satisfaction between low and high importance should, on the average higher satisfaction for facets of low importance to the workers.

A variety of facet-importance weighting schemes were applied to improve the prediction of job satisfaction from information about the amount of benefit provided by the job. None of the weighting methods provided any improvement over unweighted predictions, and some actually worsened the predictions. Possible conceptual and methodological explanations for this outcome are reviewed.

Chapter 10

JOB SATISFACTION MEASUREMENT: NORMATIVE AND INDIVIDUALISTIC MODELS

Are people better satisfied by simply having more of all the "good" features of jobs or, instead, by having more of what they personally and uniquely consider most important? While this question is discussed frequently, it is still unresolved.

It is known that there are widely-shared values that indicate for most people the facets that jointly constitute a "good job", a satisfying one. This suggests that the inducement of job satisfaction is rather uniform among employed people so that a simple summation or averaging of "good" job qualities will serve well as a predictor of the job satisfaction of the job holder. Some leading measures of job satisfaction reflect this view, the Job Description Index (Smith, Kendall and Hulin, 1969) being an example. Another is the Quality of Employment Index (Cammann, Quinn, Beehr & Gupta, Chapter 3), which is put forward as a summative index of the objective qualities of the job usable as a predictor of the satisfaction of the job occupant. However, it remains a question, both practical (how to assess or measure job satisfaction) and theoretical (from what sources does satisfaction arise) whether individual differences in preference, need, or local variations in social norms, affect the response of satisfaction sufficiently to warrant individually weighted models and measurement operations.

A number of prior investigations, some theoretical and some empirical, have addressed this problem. An early statement of a theoretical view is that of Morse (1953) who wrote, ". . . the greater amount [of some job-produced good] the individual gets the greater his satisfaction and, at the same time, the more the individual still desires, the less his satisfaction." Within the framework of expectancy theory, Vroom (1964) posited a multiplicative relationship between how much the individual gets (of some "good") from the job and how much the individual wants of it; that is, the goods and bads of the job are multiplied by their valences for the job occupant in the determination of job satisfaction. Others (e.g., Mitchell & Albright, 1972) propose that job satisfaction is a joint function of the importance of various outcomes ("goods") and the individual's belief that his job is a means for attaining such outcomes. Barth (1974) proposes that job satisfaction is a product of the individual's priorities regarding the returns from his job and his expectation that these returns will be forthcoming in the future. Lawler (1973) has developed a model that includes a specification that the increment to satisfaction arising from any facet of a job is modified by, among other things, the individual's sense of the importance of that facet to him, personally. Locke (1969) holds that the crucial factor is the discrepancy between what the individual wants or expects from a job and what he actually receives.

The empirical literature is more bland and unsatisfying than the elegant and persuasive theories. Efforts to improve the measurement of job satisfaction by using methods that incorporate individual differences have been uniformly unsuccessful (Evans, 1972; Wanous & Lawler, 1972; Quinn & Mangione, 1973; Barth, 1974). The weighting of job satisfaction indicators according to the importance or priorities reported by the individual has failed to achieve either (1) improved internal consistency and reliability of measures of job satisfaction, or (2) consistently improved predictions of behavior presumed to be influenced by the individual's job satisfaction (although there are some exceptions). Still, the use of indicators of satisfaction that do not accommodate individual differences remains incompatible with the rich evidence regarding individuality. It is hard to accept that job satisfaction is fully "explained" by the occupant's receiving standard proportions of standard benefits. A further compulsion to press this line of inquiry is that the correlations between unweighted measures of job satisfaction and the unweighted objective measures of the benefits of jobs are lower than expected; only about half of the variance in job satisfaction is statistically "explained" by unweighted job attributes. Further, the capacity of present indicators of job satisfaction to predict behavior is sometimes disappointing. and incompatible with both theory and common sense expectations.

The purpose of the investigation reported here is to advance inquiry into these issues. Specifically, the aim is to see whether job satisfaction, directly measured, is better accounted for by a normative model of causation (all people respond in about the same way to given degrees and classes of environmental return from their jobs) or by a weighted model (which assumes that people respond differently according to their own individualistic sense of importance or priority).

Method

Sample

The sample for this investigation comprised 651 employed adults in Michigan. These respondents did not, strictly speaking, constitute a sample of any defined population but were instead, a varied assortment of individuals chosen because they were in diverse jobs with five diverse employers. Included were 173 members of a printing firm, 21 members of a research and development firm engaged in research in the physical and mathematical sciences, 120 and 124 members, respectively, in two automotive supply manufacturing firms, and 213 members of the service departments of a hospital. Supervisors were included. Details of the sample are described in Chapter 1.

Measures

Four sets of individual level measures, all obtained through standardized interviews, are pertinent to this report. Common to all four are the five job facets distinguished by analysis of national survey data as a means to simplify the description of jobs with minimum loss of information. These facets are called: Resource Adequacy (equipment, information, supplies, etc.), Comfort and Convenience, Challenge (use of skills, opportunity to learn, etc.), Financial Rewards (pay, benefits, etc.), and Coworkers. Notwithstanding slight shifts in labels, a full account of the derivation of these facets and their component indicators is provided by Quinn & Shepard (1974), including specifications of their psychometric properties and evidence of validity. The four sets of measures were:

1. Facet importance. Each of the five importance indices was derived by averaging the importance ratings assigned by each respondent to each of the facet components. The general form of the questions is suggested by this example: "How desirable is it to you to have a job where the physical surroundings are pleasant?"

2. Facet amount (realization). Each index was derived by averaging the respondent's rating of the extent to which his job actually provided to him, or was characterized by, the specific beneficial component attributes. An illustration of the question format is: "To what extent is this true about your job? I have enough authority to do my best." Quinn and Shepard do not offer a realization scale for the coworkers facet.

3. Facet job satisfaction. A similar averaging process was used to derive the five indices representing the respondent's degree of satisfaction with each of the job facets. The sample question format is: "How satisfied are you with your fringe benefits?"

4. Overall job satisfaction was measured in two ways: (a) a summative average of the five facet satisfaction indices described above; and (b) an index derived by Quinn and Shepard (1974) from the following five questions containing no reference to specific job attributes.

> All in all, how satisfied would you say you are with your job--very satisfied, somewhat satisfied, not too satisfied, or not at all satisfied?
>
> In general, how well would you say that your job measures up to the sort of job you wanted when you took it? Would you say it is very much like, somewhat like, or not very much like the job you wanted when you took it?
>
> Knowing what you know now, if you had to decide all over again whether to take the job you now have, what would you decide? Would you decide without any hesi-

> tation to take the same job, would you have second thoughts, or would you decide definitely not to take the job?
>
> If a good friend of yours told you (he/she) was interested in working in a job like yours what would you tell (him/her)? Would you strongly recommend this job, would you have doubts about recommending it, or would you strongly advise (him/her) against this sort of job?
>
> If you were free to go into any type of job you wanted, what would your choice be?

The two job satisfaction measures are hereafter designated as Overall Facet Satisfaction and Facet-free Satisfaction. Table 1 shows the intercorrelations among this set of two overall job satisfaction measures, and also their correlations with the five facet importance, and four facet amount measures.

Analysis Plan

The analytic strategy involved three steps, as follows: First, a determination of whether there were differences in job satisfaction between employees rating the facets as of high or of low importance; Second, a determination of whether there were differences between those rating a facet high vs. low in importance as to the correlation between facet realization (amount) and job satisfaction; Third, a comparison of the weighted and the unweighted indicators of facet amount as to their correlations with job satisfaction

Table 1

Matrix of Intercorrelations among the Facet Importance, Facet Realization, and Overall Job Satisfaction Indexes (N=541)

	1	2	3	4	5	6	7	8	9	10
Importance										
1. Resource adequacy										
2. Comfort	$.38^{**}$									
3. Challenge	$.30^{**}$	$.20^{**}$								
4. Financial rewards	$.29^{**}$	$.47^{**}$	$.23^{**}$							
5. Coworkers	$.32^{**}$	$.45^{**}$	$.24^{**}$	$.35^{**}$						
Facet Realization										
6. Resource adequacy	$.15^{**}$	.05	$.09^{*}$	.06	$.14^{**}$					
7. Comfort	.02	-.06	.03	.01	.01	$.24^{**}$				
8. Challenge	-.03	-.06	$.28^{**}$	.07	.06	$.32^{**}$	$.24^{**}$			
9. Financial rewards	.04	-.08	$.22^{**}$	.01	.06	$.15^{**}$	$.23^{**}$	$.43^{**}$		
Overall Satisfaction										
10. Facet-free	-.01	-.03	.08	.05	.04	$.30^{**}$	$.28^{**}$	$.45^{**}$	$.23^{**}$	
11. Total facet-specific	.05	.01	$.18^{**}$	$.10^{*}$	$.15^{**}$	$.61^{**}$	$.45^{**}$	$.59^{**}$	$.34^{**}$	$.59^{**}$

* $p < 0.05$

** $p < 0.01$

Results

Table 1 shows that the measures of facet importance intercorrelate positively and significantly, as do the measures of facet realization and overall satisfaction. Table 2 demonstrates that for each facet the standard deviation (and hence the variance) of degree of facet realization exceeds that of degree of facet importance.

Relevance of Facet Importance to Job Satisfaction

One indicator that the importance of the several job attributes may influence the prediction of overall job satisfaction would be the presence of differences in satisfaction outcomes for respondents differing in their facet importance ratings. Tables 3-6 display the results of such an analysis. Table 3 shows, for example, average satisfaction scores (two measures of overall job satisfaction and one of satisfaction with Resource Adequacy) for respondents with varying degrees of Resource Adequacy (four degrees), dichotomizing the respondents on their importance ratings for Resource Adequacy.

Inspection of these tables reveals immediately a pattern of increment in job satisfaction with increments in facet realization amounts for all the facets. This is to be expected, of course, to the extent that the experience of satisfaction is derived from exposure to a more favorable job environment.

The pattern of differences between the high importance and low importance respondents is, however, more complex. The most prevalent

Table 2

Standard Deviations of Facet Importance and Facet Realization

Facet	Standard Deviation of rating on importance	Standard Deviation of rating on realization
Resource Adequacy	.67	.76
Comfort	.54	.61
Challenge	.45	.84
Financial Rewards	.47	.53

Table 3

Average Overall Job Satisfaction Scores for Groups Defined by Level of Rated Importance of Job Facet and Level of Facet Realization Amount: Resource Adequacy Facet*

Overall Job Satisfaction	Facet Importance	Facet Realization Amount			
		Low------	------	------	------High
Facet-free	Low	3.09 (18)	3.55 (44)	3.74 (79)	4.38 (9)
	High	2.99 (42)	3.14 (100)	3.78 (280)	4.09 (57)
Facet-specific	Low	2.62 (18)	2.91 (44)	3.18 (79)	3.60 (9)
	High	2.37 (42)	2.75 (99)	3.20 (280)	3.53 (57)
Satisfaction with Specific Facet					
Resource adequacy	Low	2.27 (18)	2.81 (44)	3.25 (79)	3.78 (9)
	High	2.02 (42)	2.67 (99)	3.27 (280)	3.70 (57)

* Numbers in parentheses are cell numbers.

Table 4

Average Overall Job Satisfaction Scores for Groups Defined by Level of Rated Importance of Job Facet and Level of Facet Realization Amount: Comfort Facet*

		Facet Realization Amount		
Overall Job Satisfaction	Facet Importance	Low	Medium	High
Facet-free	Low	3.22 (24)	3.55 (249)	3.88 (159)
	High	3.18 (25)	3.42 (116)	3.98 (70)
Facet-specific	Low	2.69 (24)	2.98 (249)	3.28 (160)
	High	2.46 (24)	2.98 (116)	3.34 (68)
Satisfaction with Specific Facet				
Comfort	Low	2.60 (24)	2.93 (249)	3.31 (160)
	High	2.28 (24)	2.94 (116)	3.43 (68)

* Numbers in parentheses are cell numbers.

Table 5

Average Overall Job Satisfaction Scores for Groups Defined by Level of Rated Importance of Job Facet and Level of Facet Realization Amount: Challenge Facet*

Overall Job Satisfaction	Facet Importance	Facet Realization Amount Low			High
Facet-free	Low	2.91 (9)	3.03 (79)	3.72 (150)	3.95 (128)
	High	2.50 (12)	2.85 (34)	3.30 (78)	3.95 (141)
Facet-specific	Low	2.30 (9)	2.68 (78)	3.04 (150)	3.27 (128)
	High	2.13 (12)	2.59 (34)	2.98 (77)	3.35 (141)
Satisfaction with Specific Facet					
Challenge	Low	2.07 (9)	2.47 (78)	2.95 (150)	3.36 (128)
	High	1.79 (12)	2.34 (34)	2.98 (77)	3.49 (141)

* Numbers in parentheses are cell numbers.

Table 6

Average Overall Job Satisfaction Scores for Groups Defined by Level of Rated Importance of Job Facet and Level of Facet Realization Amount: Financial Rewards Facet*

Overall Job Satisfaction	Facet Importance	Facet Realization Amount		
		Low	Medium	High
Facet-free	Low	3.07 (54)	3.65 (179)	3.83 (111)
	High	3.34 (37)	3.67 (174)	3.70 (81)
Facet-specific	Low	2.72 (53)	3.05 (179)	3.23 (112)
	High	2.59 (36)	3.11 (173)	3.19 (81)
Satisfaction with Specific Facet				
Financial rewards	Low	2.68 (53)	2.97 (179)	3.21 (112)
	High	2.58 (36)	3.06 (173)	3.30 (81)

* Numbers in parentheses are cell numbers.

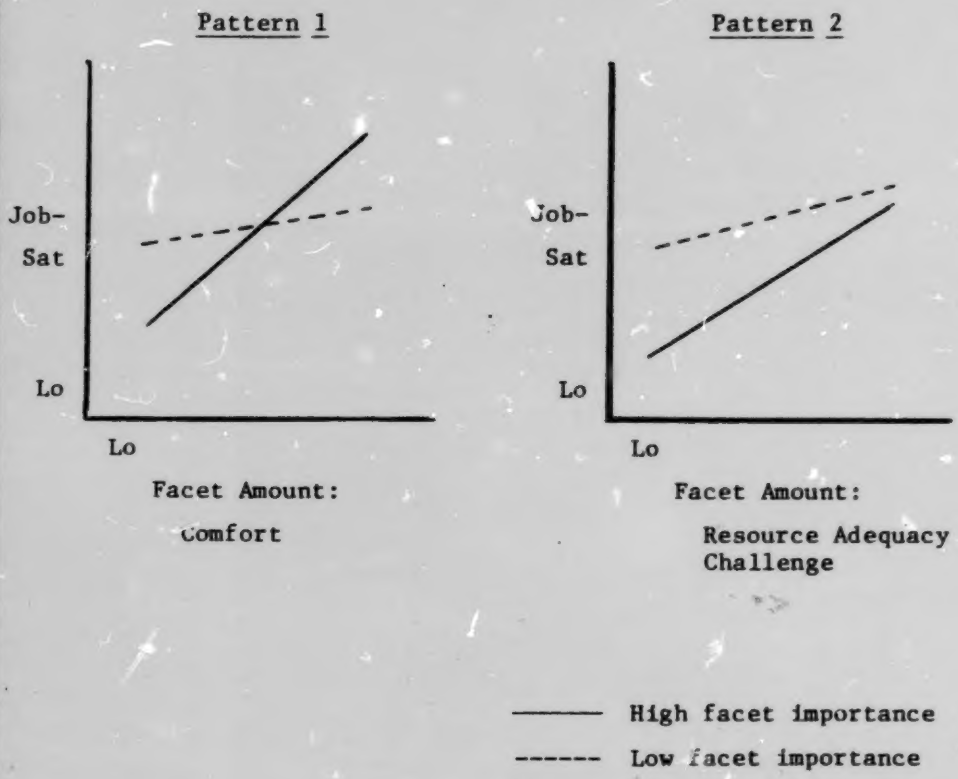
pattern, and the expected one, is for the low importance group to be less responsive than the high importance group to variations in facet realization amounts. That is, one may expect a person giving less importance to, say, Resource Adequacy to be only slightly dissatisfied with a deficit in this job quality, and to be only slightly satisfied by an abundance of it; those to whom resource adequacy is important can be expected to register greater extremes of satisfaction and dissatisfaction. This pattern holds consistently with respect to the facets of Comfort and Challenge (Tables 4 and 5) and marginally for the facets of Resource Adequacy and Financial Rewards (Tables 3 and 6).

The facets also vary somewhat on whether there is a prevailing difference in satisfaction, on average, between the low and high importance groups. Respondents who assign low importance to a facet report consistently higher satisfaction than those who assign it high importance for two of the five facets: Resource Adequacy and Challenge. A similar but more marginal trend characterizes a third facet, Comfort; and no trend emerges for the facet of Financial Rewards. The trend for greater importance of Challenge to be associated with lower satisfaction when the effects of importance are considered separately from realization effects is a reversal of the significant positive correlation observed in Table 1 for Importance-Challenge and total facet-specific satisfaction, where the effects of importance and realization are confounded. Realization evidently acts as a supressor upon the relationship of importance and satisfaction, obscuring their true negative relationship, which is, therefore, only observable when realization's effects are removed.

The diagrams in Figure 1 provide a rough summary of the patterns described above. The results of this inspection of the data suggests that individual differences in the importance of job facets might successfully be used as a weighting factor in achieving more precise measures of job satisfaction.

Figure 1

Three Patterns of Relationship between Facet Amount and Job Satisfaction for Conditions of High and Low Importance*



* No consistent pattern for Financial Rewards and the four measures of satisfaction

Correlations between Facet Realization and Job Satisfaction by Levels of Importance

The observations described above are confirmed in a more refined analysis using correlational methods. The earlier notion that the high importance group would prove more responsive to variations in facet realization also suggests higher correlations between facet realization and measures of job satisfaction among the high than the low importance group. Accordingly, Table 8 shows, separately for the high and low importance groups, the correlations between four measures of job satisfaction and the respondents' quality of employment of facet realization amount--on each of the five facets. There is plainly evident a stronger response by the high importance group than by the low importance group to the experience of deprivation or abundance, with the exception of the Financial Rewards facet which exhibits a reverse pattern.

These findings correspond to the impressions reported above regarding extremity of scores (or amount of variance). The expected pattern above on extremity of scores was closely adhered to by the facets of Comfort and Challenge, followed by the other facets. With the correlational data, facets most in compliance with predictions were again observed to be Comfort and Challenge, followed by Resource Adequacy with a reversal of the predicted pattern for Financial Rewards.

Predicting Overall Job Satisfaction from Weighted and Unweighted Predictors

There are a number of ways to compute weighted measures of job satisfaction with reference to individual differences in facet importance ratings (or in facet-specific expectations). Each has its own

Table 8

Correlations between Facet Realization Amount and Overall Job Satisfaction by Levels of Importance

Facets	Rated Importance	Facet-free	Facet-specific	JS/specific facet
Resource Adequacy	Low N=152	.21**	.52**	.68**
	High N=479	.34**	.64**	.73**
Comfort	Low N=433	.20**	.39**	.46**
	High N=208	.37**	.52**	.63**
Challenge	Low N=368	.38**	.51**	.61**
	High N=274	.50**	.64**	.72**
Financial Rewards	Low N=345	.28**	.37**	.30**
	High N=295	.11*	.31**	.30**

* $p < 0.05$
** $p < 0.01$

conceptual properties and assumptions about the processes of attitude formation. These alternative weighting schemes are treated in detail by others, and will not be described here (Wanous & Lawler, 1972; Quinn & Mangione, 1973; Barth, 1974). For the following analysis two alternative forms of weighting were used on grounds that both have shown interesting properties in other studies, and also on the practical grounds that they matched the available data for the present study.

The first form of weighting is derived from Vroom's original suggestion (1964) that net job satisfaction is the sum across facets of the importance of each facet multiplied by the amount of actual return or outcome for that facet. The essential proposition is that the impact on satisfaction of abundance or deprivation is nil for facets that are not valued at all and very great for those assigned high importance. The second form of weighting chosen is a subtractive index in which the (standardized) facet realization amounts are subtracted from the standardized importance ratings. While the conceptual foundations for the use of importance ratings in the subtractive index are problematic, such usage has in fact achieved successful results for a number of investigators (e.g., Wanous & Lawler, 1972).

The correlations between the weighted and the unweighted predictors (four facets each) on the one hand, and job satisfaction on the other are studied. Job satisfaction is represented by the two overall measures and also the respondents' specific facet satisfaction relevant in each case. The data show consistently, with but one minor exception in 24 comparisons, that weighting of job facet realization amounts by their respective degrees of importance typically has the effect of decreasing rather than increasing the concurrent prediction of job satisfaction. Multiple regressions similarly demonstrated little or no increment in prediction from the weighted predictors over the realization amounts on their own.

Discussion

This attempt to improve the concurrent prediction of job satisfaction over that available from information about the objective qualities of an individual's job and job environment was a failure. No means was found to employ information about the importance of the job facets to the individual to arrive at a more accurate concurrent prediction of his actual job satisfaction. In this failure the inquiry is not unique, as this outcome matches those of other investigators, notably Wanous & Lawler (1972), Quinn & Mangione (1973) , and Barth (1974), who used somewhat different measurement methods or data from distinctively different populations. There is still no persuasive evidence to provide support for the hypothesized superiority of importance-weighted models of job satisfaction. The conclusion reached by Quinn and Mangione can stand without modification:

> ". . . not only replicated the unanimously 'negative' findings of earlier validation studies but showed in addition that importance weighting actually decreases the validity of job satisfaction. The presumed theoretical 'help' provided by the principle of importance-weighting proved instead to be an empirical hindrance in estimating overall job satisfaction. Weighting of satisfactory ratings by importance ratings, which earlier studies had shown to be at worst a rather innocuous theoretical indulgence, was shown by the present data to be something worse than innocuous (p. 18)."

In our view, however, the matter is not yet settled. The accumulation of consistent empirical results across some variety of measurement methods, populations, and methods for weighting is not to be set aside lightly. Still, the logic of weightings is so impeccable, the intuitive appeal of the idea of individualistic response to job environments is so strong, and the presently unaccountable degree of variance among people in their experience of job satisfaction is so much in need of explanation, that one is drawn into a search for the reasons behind the discouraging results that have been reported. Also, quite apart from theory and ideology, it is statistically implausible, although not impossible, that differences in "importance" ratings such as those shown in Tables 3-6 must remain useless in improving the prediction of job satisfaction. There are two kinds of plausible reasons for failure: perhaps we fail to discover significant gains from individualistic weighting because of faulty methods; perhaps we fail because we work with an over-simplified view of how individuality is brought to bear in the context of response to jobs.

Methodological Issues

1. Choice of measures. In developing measures of job qualities and job satisfactions, investigators go to considerable trouble to narrow their attention to job facets that are important to most workers. For example, Smith, et al. (1969) and Barnowe, et al. (1972) engaged in careful preliminary work to eliminate the trivial from consideration. They may have done this work too well, in the sense that only factors or facets remain that are quite uniformly of high importance to workers. To the extent that this is the case, the individual differences obtained in importance ratings may be differences of small potential effect. This notion finds some support

in Table 2 which establishes that the variances of scores on facet realization exceed those on facet importance. Yet the difference is not so great as to suggest that scores on facet importance could have no appreciable effect as weights.

2. Scale properties of the measures. A subtractive weighting scheme is applied with risk to measures that lack comparable and equal scale intervals. Multiplicative weighting schemes are sensitive not only to the risks relevant for subtractive schemes, but also should have an absolute zero point. Perhaps, the usual steps to moderate these risks are inadequate, so that the weighting introduces more of error than of valid information. Empirical weighting schemes could be used to help deal with some of these problems--for the subtractive model by allowing the relative weights of importance and realization to be determined by a multiple regression, and the multiplicative model, by taking the log of the variables and performing a multiple regression. When this empirical approach to weighting was applied to the subtractive model for one facet (Challenge) it was found, however, that there was no increment in prediction of total facet-specific satisfaction over that available from realization alone, even though the weights generated correspond to the results suggested by Table 5, i.e. a positive weight for realization and a negative weight for importance.

3. Pre-weighting of measures. Subjective ratings of the amount of a facet in one's environment may already take into account the importance of the particular facet to a worker. This kind of preweighting of the worker's report of the amount of a facet he or she receives reduces the ability of purer measures of importance to improve prediction, as importance has already been incorporated into the measurement of working conditions. In regard to the subtractive index, however, such pre-weighting is not

substantial given the generally low correlations (see Table 1) of realization and importance.

4. High intercorrelations among predictors. In principle, a highly correlated set of predictors does not allow much room for improvement of prediction by weighting of the components. This was no problem in the present study for the subtractive index, as it is half composed of importance, which displays quite low correlations with realization ratings. It does, however, carry some weight for the multiplicative index which is more liable to be linearly related to its two components.

Conceptual Issues

1. Maximizing vs. optimizing. All of the approaches cited in this report to weighted predictors of job satisfaction share a common set of assumptions of a very simplistic sort. It is assumed that the job attributes valued by most people are in short supply such that few people have an excess, or even enough, of these returns from their job. It is, thus, a question of allocation of deficits, and the view of "optimum allocation" of deficits is plausibly to be patterned according to the individual's priorities among types or classes of job rewards. The image follows that the cause-effect relationships are linear, monotonic, and additive when weighted according to individual priorities. These assumptions are compatible with the usual methods of measurement and analysis. They are incompatible with plausible alternative conceptions. It is known that some job attributes do not have a linear and monotonic relationship with outcomes such as satisfaction (French, 1973) but display strong curvilinearity with peaks in mid-scale such that a large portion of a population may have more of that job attribute than is wanted, thus reversing the job attribute's causal relationship to satisfaction. Some theories such

as those of Herzberg, et al., (1959) and Maslow (1964), while controversial, imply the prevalence of step functions, with the "step" differently but not randomly located for different individuals. It is suggested that an exploration of the validity of the linear-monotonic assumption might disclose a means for improving our understanding of the etiology and continuing maintenance of job satisfaction.

2. Self selection and accommodation. The working American population is rather migratory with a high rate of job and employer turnover, much of it voluntary in search of a better fit between job and person. It is likely also that some people are successful in achieving a fit between the attributes of their job and their personal facet priorities by modifications in the job itself, or by the adjustment of their priorities to the realities as they come to see them. Such social and psychological processes can be expected to result in there being a sizeable proportion of employed people in jobs that have been made or chosen to "fit". In such cases, one should expect only relatively modest discrepancies between the person's felt needs and priorities on the one hand, and the attributes of the job situation on the other. For such people, the prevalent methods for determining optimum individualized weighting in the measurement of job satisfaction would be treating relatively trivial (although measureable) variations, among people who are, for the most part, highly satisfied with their jobs. A more compelling test of the weighting theory would be obtained if the work were concentrated in sub-populations selected for their vulnerability to individualized discrepancy between wants and realizations. Such populations might be those, for example, of relatively low overall job satisfaction, or those suspected of being "locked in" to their jobs, or those

who have had only a brief time for the job search and accommodation processes to have had their full effect (e.g., the new entrants to the labor force, and those relatively new in their present jobs).

3. Rare events. The contemporary approach to the issues of weighting job satisfaction predictors is designed to exploit the prevalence of common conditions as to job attributes and job satisfaction, and to fully exploit the power of large-N population statistics. One can defend this approach, but it would be wise, perhaps, to explore the possibility that essential information is missed in such an approach. It may be that the critical factors of an individualistic sort tend to be very numerous in variety with relatively few instances of each -- i.e., rare events. It is not hard to imagine (or to recall) examples of the uncommon but compelling discontinuities between job attributes and personal needs that may have the force of rendering inconsequential the other factors of fit that may be present. An approach to assess the significance, if any, of this idea could take the form of (a) modifying the measurement of job attributes and job facet priorities by including provisions for unique, or at least non-prevalent, facets and (b) to examine and compare known populations, some susceptible to standard weighting procedures and others displaying clearly the operation of some job facet not included in the standard measurements.

References

Barnowe, J. T., Mangione, T. W., & Quinn, R. P. The relative importance of job facets as incicated by an empirically derived model of job satisfaction. Ann Arbor, Mi.: Survey Research Center, 1972 (mimeo).

Barth, R. A. An empirical examination of several job attraction and job satisfaction measures. International Review of Applied Psychology, 1974, 25, 53-64.

Cammann, C., Quinn, R. P., Beehr, T. A., & Gupta, N. Effectiveness in work roles, Report I: Validation of the quality of employment indicators. See also Chapter 3.

Evans, M. G. Relations among weighted and non-weighted measures of job satisfaction. Studies in Personnel Psychology, 1972, 4, 45-54.

French, J. R. P., Jr. Person role fit. Occupational Mental Health, 1973, 3, 15-20.

Herzberg, F., Mauser, B., & Snyderman, B. B. The motivation to work. New York: Wiley, 1959.

Lawler, E. E., III. Motivation in work organizations. Monterey: Brooks/Cole, 1973.

Locke, E. A. What is job satisfaction? Organizational Behavior and Human Performance, 1969, 4, 309-336.

Maslow, A. H. Motivation and personality. New York: Harper, 1964.

Mitchell, T. R., & Albright, D. W. Expectancy theory predictions of the satisfaction, effort, performance, and retention of Naval aviation officers. Organizational Behavior and Human Performance, 1972, 8, 1-20.

Morse, N. C. Satisfaction in the white-collar job. Ann Arbor, Michigan: Survey Research Center, 1953.

Quinn, R. P., & Mangione, T. W. Evaluating weighted models of measuring job satisfaction. A Cinderella story. Organizational Behavior and Human Performance, 1973, 10, 1-23.

Quinn, R. P., & Shepard, L. J. The 1972-1973 quality of employment survey. Ann Arbor, Michigan: Institute for Social Research, 1974.

Smith, P. C., Kendall, L. M., & Hulin, C. L. The measurement of satisfaction in work and retirement. Chicago: Rand-McNally, 1969.

Vroom, V. H. Work and motivation. New York: Wiley, 1964.

Wanous, J. V., & Lawler, E. E., III. Measurement and meaning of job satisfaction. Journal of Applied Psychology, 1972, 56, 95-105.

Chapter 11

THE IMPACT OF COGNITIVE SOPHISTICATION ON RESPONSES TO AN OCCUPATIONAL SURVEY

by

Robert P. Quinn
Bernard J. Goitein
Graham L. Staines
David J. Pagnucco

ABSTRACT

Though its existence is widely acknowledged among survey researchers, the impact of cognitive sophistication on interview responses has rarely been systematically examined. This paper explores three cognitive operations in which cognitively sophisticated and cognitively unsophisticated respondents might be thought to differ. First, sophisticated respondents should display a greater capacity for abstraction than less sophisticated respondents, as evidenced by a greater disparity between their ratings of present and ideal work situations. Secondly, they should demonstrate a greater capacity for recall, as indicated by a closer correspondence between retrospectively-reported changes in working conditions and measured changes (i.e., differences between the same measures at two times). Sophisticated respondents, thirdly, should perform better at differentiation, interpreted in terms of lower correlations among measures of different dimensions of job characteristics.

Based on an occupational survey, empirical tests of the predictions regarding these three cognitive operations reveal that cognitively sophisticated respondents outperformed their less sophisticated counterparts on the indicators of abstraction and recall, but the reverse pattern was found in the case of differentiation. Possible explanations were discussed for the unexpected finding on differentiation. The implications of these results are noted for future survey research.

Chapter 11

THE IMPACT OF COGNITIVE SOPHISTICATION ON RESPONSES TO AN OCCUPATIONAL SURVEY

Survey researchers have long known that surveys elicit different answers from cognitively sophisticated and cognitively unsophisticated respondents. Sophisticated respondents give longer, more elaborate responses to open-ended questions than do their less sophisticated counterparts. They generally understand the questions better, answer "Don't know" less frequently, and are less affected in their answers by variations in question format (Schuman & Presser, 1977). On the other hand, they are more likely to get bored, to feel busy and imposed upon, to argue about the wording of certain questions, and to query the merits of the entire survey enterprise. In contrast, less cognitively sophisticated respondents are more inclined to act as if being given a test, to worry that their answers are wrong, to apologize for their lack of knowledge, to try to give the answer they think the interviewer wants, and to terminate an interview prematurely when it is overly abstract or confusing.

Designers of surveys try to counteract some of the effects of respondents' sophistication. For example, questions are generally framed in language that virtually all respondents can understand. Interviews are kept short enough to maintain the attention of the less sophisticated and the cooperation of the more sophisticated. Alternatively, modes of measurement other than self-report may be adopted to offset the method variance associated with personal interviews.

One of the more thoughtful analyses of the effects of respondent sophistication appears in the political literature. Converse (1964) cites persuasive evidence that a wide variety of attitudinal and ideological measures on political topics will intercorrelate more highly among politically sophisticated than among politically unsophisticated respondents. His point is that there is greater intellectual coherence or constraint among the cognitions of sophisticated people, and that such internal consistency translates into higher correlations. Bishop (1976) adds that formal education provides much of the intellectual sophistication needed for such consistency among political attitudes.

The present paper attempts to extend the views of Converse and Bishop to a more differentiated set of predictions of responses to an occupational survey. Under some circumstances, that is, correlations between measures will be greater among more sophisticated respondents, but not always. Specifically, this paper proposes three cognitive operations on which intellectually sophisticated respondents would be expected to outperform those less sophisticated: abstraction, recall and differentiation. Each of the three cognitive operations generates a specific prediction.

Abstraction

The capacity for abstract thinking is one of the core dimensions of intellectual performance. Accordingly, cognitively sophisticated respondents should display a superior capacity to entertain hypothetical possibilities, including ideal situations that diverge from their actual circumstances. In this connection, Moch, Cammann and Cooke (1976) obtained ratings of actual influence and desired influence in various occupational domains from respondents who differed in terms of education, a useful proxy for cognitive sophistication. Specifically, they collected data from respondents in a manufacturing firm (low education, mainly high school) and an engineering organization (high education, mainly college graduates) concerning their influence in three domains: work activities, activities involving the allocation of personnel resources, and coordination activities. Moch et al. found the respondents in the manufacturing firm (low education) much less able to distinguish between actual and desired influence over personnel resources and over coordination activities than their counterparts in the engineering organization (high education). No such difference emerged between the groups with respect to actual and desired influence over work activities.

Ratings of the importance of job facets offer another, albeit more indirect, means of tapping the ability to abstract to hypothetical situations. When respondents cite a set of job facets as important to them, it may be inferred that their ideal jobs would score high on these facets. Thus, importance ratings may be interpreted as indicating the characteristics of ideal jobs, as envisaged by respondents. Cognitively sophisticated respondents should be able to draw a sharper distinction between the jobs they currently hold and the types of jobs they would like to have, as indicated by their ratings of facet importance. The prediction implicit here may be stated in operational terms as hypothesis one.

Hypothesis one. The correlations between the descriptive ratings of job facets on the present job and the ratings of the importance of the same job facets should be higher among the less cognitively sophisticated respondents than among the more sophisticated ones.

Recall

The ability to recall points to a second point of divergence between more and less cognitively sophisticated respondents. Sophisticated workers should provide more accurate retrospective estimates of job change. The comparison implicit here is between retrospectively-estimated change and measured change, the latter defined as the discrepancy between ratings of job facets provided by the same workers on two separate occasions. Hypothesis two, thus, assumes the following operational form:

Hypothesis two. The relationship between retrospective estimates of change and measured change should be stronger among the more than among the less cognitively sophisticated respondents.

One relevant study (U.S. Department of Health, Education, and Welfare, 1965) examined the accuracy of reports of previous hospital episodes (excluding maternal deliveries). It measured accuracy of recall in terms of the absence of retrospective underreporting of the number of one's hospital episodes over the preceding year. The investigators reported a weak, positive, but not entirely linear relationship between education and accuracy of retrospective recall in the original interviews. Additional data from follow-up interviews established a weak, positive, and predominantly linear association between education and two measures of accuracy of recall: (1) correct reporting in both interviews (scored in a positive direction), and (2) failure to report in both interviews (scored in a negative direction).

Differentiation

Cognitively sophisticated respondents may be presumed to perceive their environments as highly complex and should, then, evidence a greater capacity to differentiate among various job facets. Thus, given a set of measures that tap different dimensions of task characteristics, a sophisticated respondent should provide a relatively differentiated pattern of responses. In comparison, a less sophisticated worker should offer a more homogeneous set of ratings. In accordance with the notion of a halo effect, the worker should tend to rate the job as favorable on all dimensions, or unfavorable in all respects. Hypothesis three summarizes the major prediction in operational terms.

Hypothesis three. The correlations among ratings of different job facets should be higher among the less than the more cognitively sophisticated respondents.

Some data relevant to this hypothesis were obtained in the study by Moch et al. (1976). As noted earlier, these investigators asked respondents in two organizations to rate the extent to which they felt they exercised influence in work activities, activities involving the allocation of personnel resources, and coordination activities. Moch et al. reported the degree of relationship between the three influence domains for respondents in each organization. Contrary to hypothesis three, for two of the three pairings of influence domains (work-personnel and work-coordination), the inter-domain correlations among the engineers exceeded those among the employees of the manufacturing firm.

As for studies more marginally relevant to hypothesis three, several authors have argued that field independence, a concept analogous to and perhaps partially overlapping with cognitive sophistication (Vernon, 1972), implies a greater capacity to differentiate. Gruenfeld and Arbuthnot (1969) found some evidence that, on a task which required subjects to rate ten peers on three scales, high scorers on field independence distinguished more sharply among the ten individuals rated, and among the three rating scales used, than did low scorers. The concept of "cognitive complexity" likewise implies a greater capacity to differentiate among dimensions in person perception. Schneier (1977) used Bieri's measure of cognitive complexity (Bieri, Atkins, Briar, Leaman, Miller & Tripodi, 1966) in an experimental study and found that for two kinds of rating scales, the more complex subjects displayed weaker halo effects in their perceptions of others than did the less complex subjects. Bieri et al.'s (1966) measure correlated only marginally with such scholastic aptitude measures as the SCAT scales (0.15 and 0.19, respectively, for SCAT-V and SCAT-Q; Vannoy, 1965) and correlated 0.01 with ratings of "academic intelligence" (Sechrest & Jackson, 1961). In line with Bieri's (1961) conclusion based on an earlier review of the evidence, the more recent studies cited above demonstrate little overlap between his measure of cognitive complexity and intelligence measures.

Using a single data set, the present study endeavors to operationalize and then test the three hypotheses listed above. The measures of cognitive sophistication available in this study are intelligence and education.

Method

Sample

The overall design of this two phase study, of which this investigation is a part, is described in detail in Chapters 1 and 2. The present report analyzes data from the 651 respondents who participated in the first phase of the study and the 272 respondents involved in the second wave of data gathering about 20 months later.

The 651 respondents in Phase I belonged to five midwest organizations (a printing firm, a research and development firm, two automotive supply companies, and a hospital). The 272 respondents interviewed in Phase II, however, belonged to only the latter three organizations. Of the 651 workers interviewed in Phase I, 579 were also observed at their jobs by trained observers who rated the jobs on a diverse set of job characteristics. The training of observers and the observational procedures in Phase I are described in detail by Jenkins, Nadler, Lawler, and Cammann (1975), and reproduced in Chapter 4 of this report. Of the 272 Phase II respondents interviewed, 147 were also observed at their jobs by a new and differently trained set of observers. The modifications in the training of observers and the procedures for observing in Phase II are described by Jenkins and Nadler in Chapter 7.

All 272 respondents in Phase II belonged to one of three subcategories according to their job status in Phase II relative to Phase I:

1. One hundred and sixty three Phase II respondents who were in the same jobs they had occupied in Phase I.
2. Seventy three Phase II respondents who were in different jobs from the ones they had held in Phase I.

3. Thirty six Phase II respondents who replaced some Phase I respondents who had left their original jobs between phases.

More detailed discussion of the sample composition appears in Chapters 1 and 2.

Measures

Dependent Variables. Thirty one items concerning "good" versus "bad" working conditions comprised the Quality of Employment (QoE) facets' scales (Challenge, Comfort, Financial Rewards, and Resource Adequacy--Barnowe, Mangione, & Quinn, 1972) and were included in the interview schedules used in both phases of the study. Accordingly, the separate QoE facets' scores were computed for each respondent, as were the total of all 31 items. The present QoE scales contain only 31 of the 33 items used by Barnowe et al. (1972). Omitted from the present study were two of the 11 items in the QoE scale for Comfort. The correlation between the nine item and 11 item scales, nonetheless, is 0.93.

Shortened forms of the four importance scales corresponding to the four QoE facets (i.e., Challenge, Comfort, Financial Rewards, and Resource Adequacy--Quinn & Cobb, 1971) were administered and mean scores computed for both phases. A mean of the four facets' scores was also computed to obtain an Importance/Total measure.

To measure retrospectively-reported change, a question was asked in the interview at Phase II, with regard to each of 12 aspects of the job: "During the last 18 months, has there been a major change for the worse, no major change, or a major change for the better on your job regarding ...?" Ten of these aspects were grouped according to the four QoE facets, and mean scores on reported change were computed for each facet. The mean of all ten aspects was also computed to generate a facet-specific measure of overall reported change.

A single facet-free measure of overall reported change was, likewise, available. It compared respondents' jobs at Phase 1 and Phase II on a scale with five verbal response categories that ranged from "a lot worse" to "a lot better."

Interview measures were constructed for five dimensions of intrinsically rewarding job characteristics. These dimensions included the four core dimensions identified by Hackman and Lawler (1971), labeled here as autonomy, variety, task completeness, and task feedback; and a dimension of required skills suggested by Jenkins et al. (Chapter 4). In order to operationalize the five dimensions, twelve interview items were selected as Phase I measures, and fifteen as Phase II measures. For most dimensions, some of the items used were not identical in the two phases.

Separate observational measures of these five dimensions were constructed and reported for each phase (Jenkins et al., Chapter 4;[1] Jenkins & Nadler, Chapter 7). The observational training procedures and measures also changed somewhat between phases of the study. For each phase and each dimension, the mean of the relevant items was computed separately for the interview measure and for the observational measure.

Control Variables (Cognitive Sophistication). Interviewers' ratings of respondents' general intelligence were included in each Phase using a five-point scale ranging from "very low" to "very high". Data from pretest samples indicated that these ratings correlated 0.72 with the Amons Quick Test of intelligence. The autocorrelation of the ratings of intelligence (with different raters over a 20 month lag) was 0.46. The scores from both phases were, accordingly, summed to provide a more reliable scale of rated intelligence for the 236 respondents who were interviewed in both phases.

---

[1] The dimensions of task completeness and required skills are referred to by Jenkins et al. as task identity, and skills and abilities, respectively.

Since number of years of formal schooling was the other control variable used to measure cognitive sophistication, the correlation obtained between rated intelligence and education was relatively encouraging (r = 0.61).

Reliabilities. Measures of internal consistency tend to give relatively unbiased estimates of a scale's reliability only for homogeneous scales. The QoE scales were not designed as homogeneous measures. Accordingly, only lower bounds for the reliability of the QoE scales could be derived from the alphas computed; these are reported in Table 1. Another source of lower bounds for the reliability of the QoE scales was the autocorrelations of the scales over the 20 month lag in the present study, among the 163 respondents in the same jobs in both phases. The autocorrelations appear in column two of Table 1. It was hoped that the higher of the two lower bounds for each QoE facet that appear in column three of the table would afford an estimate not too far below the true reliability of the facet in question.

In Chapter 8, Goitein used a path analytic approach to measurement error to assess lower bounds for the reliabilities of the QoE facets. He identified a lower bound for the reliability of the full 33 item measure of QoE/Total of 0.66, which does not differ from the 0.66 found for the present 31 item scale. Again using path analytic procedures, Goitein reported the reliability of a particular linear combination of the QoE scales used by Barnowe et al. (1972) as 0.70. The reliabilities

Table 1

Lower Bound Reliability Estimates of the Quality of Employment Facets

QoE Facet	Estimate of Lower Bound: Alpha[a]	Estimate of Lower Bound: Autocorrelation[b]	Higher estimate of the Lower Bound
Challenge	0.72	0.72	0.72
Comfort	0.39[c]	0.46	0.46
Resource Adequacy	0.51	0.47	0.51
Financial Rewards	0.12	0.61	0.61
QoE/Total	--[d]	0.66	0.66

[a] Alphas computed on the 1,497 respondents of the 1972-73 Quality of Employment Survey (Quinn & Shepard, 1974).

[b] Autocorrelations computed among the 163 respondents of the present study, who remained in the same jobs over the 20 months intervening between Phase I and Phase II.

[c] By comparison to this nine item scale, the alpha for the 11 item scale is 0.56.

[d] QoE/Total is so clearly heterogeneous that a homogeneity estimate is completely unsuitable as a measure of reliability.

in column three in Table 1 were combined using Nunnally's (1967, p. 231) formula for the reliability of linear combinations, and an estimated reliability of 0.74 was found for this same linear combination of the present scales. Such convergence among the findings obtained from different approaches to estimating the reliability of the QoE facet scales is encouraging and inspires confidence in the accuracy of these lower bound estimates.

Measures of internal consistency of the shortened forms of the Importance facets' scales are shown in Table 2. They were not particularly high (0.54 to 0.60), indicating that shortened forms of the Importance scales may not have been desirable substitutes for the original scales.

The alphas computed for the subscales on retrospectively-reported change were only marginally adequate, except for Reported Change/Challenge (0.80, Table 2). The quite acceptable alpha computed for Reported Change/Total, 0.78, probably underestimated the measure's reliability because the scale measured change in diverse aspects of the job.

For each phase, Jenkins and his colleagues (Chapters 4 and 7) computed several forms of reliability coefficients for each of the observational measures of the intrinsically rewarding dimensions of the job (autonomy, variety, task completeness, task feedback, and required skills). The alpha coefficients of these scales appear in Table 1. Alpha coefficients were also computed for the interview measures of these scales and appear in Table 3. The table demonstrates adequate reliabilities in both phases for all the observational scales and for two of the five interview scales as well. Reliabilities of the interview measures were marginal for task completeness (0.56 and 0.53 for Phases I and II, respectively) and for variety at Phase II (0.61), and not computable for the single item measure of task feedback.

Table 2

Internal Consistency Coefficients and Numbers of Items in Importance and Retrospective Change Scales

Measures	No. of Items	Reliability
Importance:[a]		
Challenge	4	0.60
Comfort	3	0.54
Financial Rewards	3	0.58
Resource Adequacy	1	--[c]
Retrospectively-reported change:[b]		
Challenge	2	0.80
Comfort	3	0.46
Financial Rewards	3	0.52
Resource Adequacy	2	0.62
Total	10	0.78

[a] Alphas computed on all 651 Phase I respondents.

[b] Alphas computed on all 272 Phase II respondents.

[c] Single item scale.

Table 3

Internal Consistency Measures and A Number of Items in Interview and Observational Measures of Five Intrinsically Rewarding Dimensions of Work

	Phase I				Phase II			
	Interview[b]		Observation[c]		Interview[d]		Observation[d]	
Dimension	Alpha	No. Items	Alpha	No. Items	Alpha	No. Items	Alpha	No. Items
Variety	.77	4	.96	3	.61	2	.93	3
Autonomy	.72	3	.96	4	.86	7	.95	5
Task completeness	.56	2	.82	3	.53[e]	2	.87	2
Task Feedback	--[a]	1	.85	3	--[a]	1	.84	2
Required Skills	.79	2	.98	4	.76	3	.91	3

[a] Single item scale.

[b] Alpha computed on all 651 Phase I respondents.

[c] Alpha reproduced from Jenkins et al. report in Chapter 4.

[d] Alpha reproduced from Jenkins & Nadler's report in Chapter 7.

[e] Personal communication from D. Jenkins, June 1977.

Analysis Plan

Hypothesis one. The test of hypothesis one required correlations between quality of employment (actual job) and the importance ratings (ideal job) on the overall measures of the job, as well as on the four specific facets (Challenge, Comfort, Financial Rewards, and Resource Adequacy). The correlations were calculated and presented separately for two levels of rated intelligence, below the median and above the median; and, likewise, for three levels of education, less than a high school diploma, high school diploma, and more than a high school diploma.

On the issue of sample size, the analysis samples used to test hypothesis one in Phase I included, for levels of intelligence, those interviewed at both phases (N = 236) and, for the control on education, all respondents interviewed in Phase I (N = 651). For Phase II, the analysis sample for intelligence consisted of all respondents interviewed at both phases (N = 236), and for levels of education all respondents interviewed in Phase II (N = 272).

Hypothesis two. Hypothesis two was tested using two statistical procedures, the first of which involved correlations between retrospectively-reported change and measured change. As used here, measured change consisted of the Phase II QoE facet score minus the corresponding score for Phase I. While it would be appropriate to correct the measured change scores for the reliabilities of the measures of each phase (Lord, 1963), this was not possible since unbiased reliability estimates were not available for the QoE facets.

The correlations between reported and measured change were presented for two measures of overall reported change (facet-specific and facet-free) against a single facet-specific indicator of measured change, and for measures of reported and measured change on the four specific facets. The correlations were computed for the two levels of intelligence and three levels of education cited above.

In addition, the correlations were presented separately for the subgroup of workers who changed jobs between Phases I and II. In all likelihood, this group experienced more change in job characteristics and, therefore, offered a more compelling test of the hypothesis. When split into three levels of education, however, the group of job changers produced cell sizes as low as 19, thus making significant differences more difficult to obtain.

The predicted differences between groups in the sizes of correlations could, however, have arisen from occupational differences, such that there were fewer changes in working conditions on the jobs of the less sophisticated respondents and, thus, a restriction of range that reduced the correlations obtained among the less sophisticated respondents. Raw regression coefficients (or b weights) were, accordingly, used to predict retrospectively-reported change on the QoE facets from measured change on these same facets, within levels of the two control variables. These regression coefficients are insensitive to problems of restriction of range and indicate the mean impact in number of scale points (of the dependent variable) of a difference of one scale point of the independent variable. Significance tests of differences among regression weights were performed using the procedures developed in the framework of analysis of covariance for the homogeneity of slopes for different groups.

The sample sizes involved in the tests of hypothesis two were restricted for both intelligence and education to the number of respondents interviewed at both phases (N = 236). For an additional analysis within one subgroup, the sample size was further restricted to the number of workers who changed jobs between the two phases (N = 73).

Hypothesis three. To test hypothesis three, intercorrelations were computed among the measures of the five task dimensions: variety, autonomy, task completeness, task feedback, and required skills. These correlations were computed for each of the levels of rated intelligence and education referred to above and for each phase of data collection. These correlations were obtained separately for Phases I and II. Hypothesis three predicts higher correlations among those low than those high on intelligence because of their limited capacity to discriminate, and, similarly, higher correlations among those in the low level of education (less than a high school diploma) than those in the high level of education (more than a high school diploma).

Of course, the predicted correlational pattern could have arisen for an entirely different reason, namely, that in objective terms the five task dimensions were more closely intercorrelated in the jobs held by the less as opposed to the more intellectually sophisticated. This is possible since the more (or less) sophisticated workers tend to be concentrated in different types of jobs. To address this alternative interpretation, a set of correlations parallel to those cited above was derived from the ratings of observers. Data from respondents and observers were then compared as follows: Observers were expected to provide relatively objective estimates of the correlations among different task dimensions for different jobs. Were respondents to display the predicted pattern of correlational differences to a greater extent than observers, hypothesis three would be supported.

The different analysis samples used to test hypothesis three varied widely in size and are therefore potentially confusing. The samples for intelligence were restricted to those workers interviewed in both phases (since the measure of intelligence was based on two separate ratings by interviewers) and, further, to those observed in a given phase (N = 210 and N = 116 for Phases I and II, respectively). For the control on education, the samples were restricted merely to those interviewed and observed in a given phase (N = 379 and N = 147 for Phases I and II, respectively).

Results

Hypothesis One: Actual and Ideal Job

The data in Tables 4-7 test hypothesis one's assertion that cognitively sophisticated workers should distinguish more sharply than do less sophisticated workers between the jobs they currently hold and the jobs they would like to hold. Specifically, the tables indicate that, as predicted, the correlations between QoE (representing the jobs workers have) and importance (representing the jobs workers would like to have) tend to be higher for the less than for the more sophisticated respondents. In Phase I, workers rated low on intelligence generate a higher correlation (scored in a positive direction) than those rated high between the overall indices of Quality of Employment and Importance (Table 4). The difference

Table 4

Phase I Correlations between Quality of Employment and Importance Ratings, by Facets and by Levels of Rated Intelligence
(N's appear in parentheses)

QoE and Importance Rating	Low Intelligence	High Intelligence
Specific Facet		
Challenge	.24* (108)	.21* (115)
Comfort	-.06 (107)	-.10 (115)
Financial Rewards	.12 (107)	.02 (114)
Resource Adequacy	.21* (104)	.10 (114)
Overall	.15 (108)	.05 (115)

*$p < 0.05$
**$p < 0.01$

between the correlations, nonetheless, does not attain statistical significance. A similar pattern of insignificant differences in the direction predicted characterizes the correlations for the four specific facets. In Table 5, the data on education in Phase I tell a similar story. For the overall measures and for the four specific facets, the correlations for respondents with less than a high school diploma are more positive than the correlations for those whose education goes beyond a high school diploma, though none of the differences is statistically significant.

In Phase II the results are broadly similar and, thus, again in line with hypothesis one. On intelligence, the correlations in Table 6 follow the expected pattern (higher for low intelligence) for the overall measures and for all specific facets except Resource Adequacy. In addition, the differences between the correlations for high and low intelligence are significant for the overall indices and for Comfort. The data on education in Phase II support hypothesis one for the overall ratings of the job and for all specific facets except Resource Adequacy (Table 7). In the cases of Comfort and Financial Rewards, the correlations for the two extreme high and low education groups differ significantly.

Hypothesis Two: Reported versus Measured Change

Shown in Tables 8-11 are the data concerning hypothesis two. This hypothesis asserts that cognitively sophisticated workers should do better than those less sophisticated at recalling accurately and reporting retrospectively on the changes in their jobs between Phases I and II. Table 8 presents the correlations between change reported retrospectively and measured change (i.e., Phase II minus Phase I scores), by high and low intelligence. As regards overall change, the correlations are appreciably but

Table 5

Phase I Correlations between Quality of Employment and Importance Ratings, by Facets and by Levels of Education (N's appear in parentheses)

QoE and Importance Ratings [a]	Less Than High School Diploma	High School Diploma	More Than High School Diploma
Specific Facet			
Challenge	.23** (171)	.35** (269)	.20** (203)
Comfort	-.03 (171)	-.05 (268)	-.09 (202)
Financial Rewards	.09 (168)	.10 (266)	.03 (203)
Resource Adequacy	.21** (166)	.15* (265)	.09 (200)
Overall	.14 (171)	.17** (269)	-.02 (203)

*$p < 0.05$
**$p < 0.01$

Table 6

Phase II Correlations between Quality of Employment and Importance Ratings, by Facets and by Levels of Rated Intelligence
(N's appear in parentheses)

QoE and Importance Ratings	Low Intelligence	High Intelligence
Specific Facet		
Challenge	.43** (104)	.28** (115)
Comfort[a]	.28** (104)	-.08 (113)
Financial Rewards	.05 (103)	.00 (113)
Resource Adequacy	.12 (102)	.19* (112)
Overall[b]	.35** (104)	.07 (115)

[a] Difference of correlations between groups is significant, $p < 0.01$.

[b] Difference of correlations between groups is significant, $p < 0.03$.

*$p < 0.05$
**$p < 0.01$

Table 7

Phase II Correlations between Quality of Employment and Importance Ratings, by Facets and by Levels of Education
(N's appear in parentheses)

QoE and Importance Ratings	Less Than High School Diploma	High School Diploma	More Than High School Diploma
Specific Facet			
Challenge	.44** (76)	.26* (92)	.43** (92)
Comfort[a]	.28* (77)	.01 (90)	-.02 (90)
Financial Rewards[b]	.29** (81)	-.02 (83)	-.29** (91)
Resource Adequacy	.14 (75)	.11 (90)	.19 (90)
Overall	.34** (76)	.13 (92)	.19 (92)

[a] Difference of correlations between less than high school diploma group and more than high school diploma group is significant, $p < 0.05$.

[b] Difference of correlations between less than high school diploma group and more than high school diploma group is significant, $p < 0.0002$.

*$p < 0.05$
**$p < 0.01$

Table 8

Correlations and Raw Regression Coefficients of Reported Change on Measured Change, by Facets and by Levels of Rated Intelligence (N's appear in parentheses)

Change Variable	r		b	
	Low Intelligence	High Intelligence	Low Intelligence	High Intelligence
Specific Facet				
Challenge[a]	.10 (103)	.45** (114)	.07 (103)	.41 (114)
Comfort[b]	.25* (103)	-.08 (112)	.17 (103)	-.05 (112)
Financial Rewards	.09 (101)	.13 (111)	.06 (101)	.07 (111)
Resource Adequacy	.37** (100)	.34** (110)	.23 (100)	.21 (110)
Overall				
Facet-Specific measure of Reported Change and Facet-Specific measure of Measured Change	.23* (100)	.35** (111)	.17 (100)	.30 (111)
Facet-Free measure of Reported Change and Facet-Specific measure of Measured Change	.25* (103)	.40** (114)	.59 (103)	1.08 (114)

---

[a] r's significantly different, r to z, $p < 0.01$; b's significantly different, $F_{1,213} = 10.50$, $p < 0.002$.

[b] r's significantly different, r to z, $p < 0.02$; b's significantly different, $F_{1,211} = 5.99$, $p < 0.02$.

*$p < 0.05$
**$p < 0.01$

not significantly larger for high than for low intelligence on both the facet-specific and facet-free retrospective measures of overall reported change. As for the individual facets considered in the same table, the predicted pattern holds for Challenge (significant difference) and Financial Rewards, but not for Comfort (significant difference in the reverse direction) or Resource Adequacy. Table 9 depicts a highly similar correlational pattern for the two extreme levels of education. In fact, comparisons between the high and low levels of education reveal only two departures from the pattern noted above for intelligence. First, the trend in the predicted direction for the facet-free measures of overall reported change proves statistically significant; and, second, the trend in the reverse direction for Comfort falls short of significance.

Despite the inevitable reduction in sample size involved, hypothesis two receives a more convincing test among the subsample of respondents who changed jobs between Phases I and II and who, thereby, experienced more changes to recall and report accurately. As Table 10 demonstrates, the discrepancies between the correlations in high and low categories of intelligence are generally larger for the subsample of respondents who changed jobs. This trend within pairs of correlations toward greater differentiation in the predicted direction among job changers applies to the facet-specific measure of overall reported change, as well as to Challenge (difference still significant in the expected direction) and Financial Rewards, with the correlations for Comfort no longer significantly different in the reverse direction. The parallel data on job changers by levels of education in Table 11 are plagued by still smaller sample sizes, yet prove quite favorable to hypothesis two. For the facet-specific

Table 9

Correlations and Raw Regression Coefficients of Reported Change on Measured Change, by Facet and by Levels of Education (N's appear in parentheses)

Dimension Pair	r			b		
	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)
Specific Facet						
Challenge[a]	-.02 (66)	.43** (79)	.32** (78)	-.02 (66)	.36 (79)	.27 (78)
Comfort	.15 (67)	.09 (77)	.06 (77)	.09 (67)	.06 (77)	.04 (77)
Financial Rewards	.02 (70)	.12 (70)	.20 (77)	.01 (70)	.09 (70)	.11 (77)
Resource Adequacy	.44** (64)	.34** (78)	.33** (74)	.26 (64)	.21 (78)	.23 (74)
Overall						
Facet-Specific measure of Reported Change and Facet-Specific measure of Measured Change	.14 (66)	.36** (75)	.33** (76)	.12 (66)	.28 (75)	.27 (76)
Facet-Free measure[b] of Reported Change and Facet-Specific measure of Measured Change	.11 (66)	.31** (79)	.48** (78)	.31 (66)	.71 (79)	1.32 (78)

[a] r's of extreme education groups significantly different, r to z, $p < 0.05$; b's differ significantly in all three groups, $F_{2,217} = 4.07$, $p < 0.02$.

[b] r's of extreme education groups significantly different, r to z, $p < 0.02$.

*$p < 0.05$
**$p < 0.01$

Table 10

Correlations and Raw Regression Coefficients of Reported Change and Measured Change among Workers who Changed Jobs, by Facets and by Levels of Rated Intelligence
(N's appear in parentheses)

Change Variable	r		b	
	Low Intelligence	High Intelligence	Low Intelligence	High Intelligence
Specific Facet				
Challenge[a]	.16 (32)	.63** (39)	.11 (32)	.52 (39)
Comfort	.15 (32)	.03 (39)	.12 (32)	.03 (39)
Financial Rewards	.12 (29)	.26 (38)	.06 (29)	.21 (38)
Resource Adequacy	.61** (31)	.52** (37)	.34 (31)	.28 (37)
Overall				
Facet-Specific measures of Reported Change and Facet-Specific measures of Measured Change	.36 (31)	.59** (38)	.21 (31)	.50 (38)
Facet-Free measures of Reported Change and Facet-Specific measures of Measured Change	.21 (32)	.34* (39)	.50 (32)	1.00 (39)

---

[a] r's significantly different, r to z, $p < 0.025$; b's significantly different, $F_{1,67} = 6.24$, $p < 0.02$.

* $p < 0.05$

** $p < 0.01$

Table 11

Correlations and Raw Regression Coefficients of Reported Change on Measured Change among Workers who Changed Jobs, by Facets and by Levels of Education (N's appear in parentheses)

Change Variable	r: Less Than High School Diploma (Low)	r: High School Diploma (Medium)	r: More Than High School Diploma (High)	b: Less Than High School Diploma (Low)	b: High School Diploma (Medium)	b: More Than High School Diploma (High)
Specific Facet						
Challenge	.02 (23)	.68** (23)	.39* (26)	.01 (23)	.45 (23)	.41 (26)
Comfort	.17 (23)	.02 (23)	.19 (26)	.14 (23)	.02 (23)	.18 (26)
Financial Rewards	.00 (23)	.16 (19)	.41* (26)	.00 (23)	.12 (19)	.25 (26)
Resource Adequacy	.54** (22)	.64** (23)	.50** (24)	.24 (22)	.46 (23)	.25 (24)
Overall						
Facet-Specific measures of Reported Change and Facet-Specific measures of Measured Change	.08 (23)	.62** (22)	.50** (25)	.05 (23)	.40 (22)	.45 (25)
Facet-Free measures of Reported Change and Facet-Specific measures of Measured Change	.09 (23)	.32 (23)	.45* (26)	.27 (23)	.60 (23)	1.54 (26)

*$p < 0.05$
**$p < 0.01$

measures of overall reported change, the comparisons between high and low education exhibit sharper differences in the predicted direction among respondents who changed jobs. The same pattern holds for Financial Rewards. For Comfort, the correlation is found here to be larger for high than for low education, though a trend in the reverse direction may be observed for Resource Adequacy.

There remains, nonetheless, the previously mentioned possibility that restriction of range in the indices of measured change in the jobs of the less sophisticated workers may explain the reduced correlations in that group. Raw regression coefficients, which are insensitive to restriction of range, were, therefore, computed in addition to the correlations (Tables 8-11). As regards the b weights, the directions of differences between groups virtually replicate those reported earlier for the correlations. Similarly, the same groups and variables that display significant differences for correlations also display significant differences for the b's, with one exception for the larger group that includes both those who did and those who did not change jobs. Whereas a significant correlational difference between extreme educational groups appears for the facet-free measure of overall reported change (Table 9), only marginally significant differences emerge among the three educational groups as tested by an omnibus F-test for homogeneity of slopes of the three groups ($F_{2,217} = 2.80$; $p < 0.06$). It should be noted that significance tests for the correlation coefficients compare only the two extreme educational groups, while the regression weights are tested over all three groups.

Hypothesis Three: Dimensions of Task Characteristics

Tables 12-15 summarize the data relevant to hypothesis three. According to the hypothesis, task characteristics should intercorrelate more highly (in a positive direction) among the less than among the more cognitively sophisticated workers, because less sophisticated workers are hypothesized to have less capacity to discriminate. Overall, the pattern observed in the tables dramatically contradicts this hypothesis. The data demonstrate clearly that the more sophisticated respondents display higher intercorrelations on ratings of job facets than do those who are less sophisticated. The halo effect, in short, belongs to the more, rather than to the less, sophisticated individuals. Since the task characteristics are so confounded with one another, the results of multiple tests of statistical significance would be misleading. Accordingly, only the overall pattern of results is reported, and significance tests of the differences were not performed.

In Phase I, this unexpected pattern occurs for intelligence in seven of the ten correlational pairs based on the five task dimensions (Table 12). For the remaining three pairs (variety/autonomy, variety/task feedback, and task completeness/task feedback), the correlations are identical for high and low intelligence. The data on education in Phase I convey a similar impression (Table 13). The comparisons between high education (more than a high school diploma) and low education (less than a high school diploma) violate the predicted pattern in nine cases out of ten, and in the tenth case, variety/task feedback, the correlations for high and low education are again identical (Table 13)

Table 12

Multimethod Phase I Correlations for Five Job Dimensions, by Rated Intelligence (N's appear in parentheses)

Dimension Pair	Interview		Observation	
	Low Intelligence	High Intelligence	Low Intelligence	High Intelligence
Variety				
/Autonomy	.51** (98)	.51** (110)	.88** (100)	.93** (110)
/Task Completeness	.06 (98)	.25** (110)	.53** (100)	.67** (110)
/Task Feedback	.17 (99)	.17 (110)	.33** (100)	.32** (110)
/Required Skills	.71** (99)	.75** (110)	.83** (100)	.86** (110)
Autonomy				
/Task Completeness	.10 (98)	.32** (110)	.58** (100)	.63** (110)
/Task Feedback	.14 (98)	.32** (110)	.33** (100)	.26** (110)
/Required Skills	.38** (98)	.45** (110)	.84** (100)	.84** (110)
Task Completeness				
/Task Feedback	.21* (98)	.21* (110)	.61** (100)	.55** (110)
/Required Skills	.11 (98)	.24* (110)	.45** (100)	.70** (110)
Task Feedback				
/Required Skills	.13 (99)	.30** (110)	.34** (100)	.36** (110)

*$p < 0.05$
**$p < 0.01$

Table 13

Multimethod Phase I Correlations for Five Job Dimensions, by Education
(N's appear in parentheses)

	Interview			Observation		
Dimension Pair	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)
Variety						
/Autonomy	.48** (144)	.56** (235)	.59** (195)	.90** (146)	.88** (235)	.87** (195)
/Task Completeness	.14 (143)	.18** (235)	.30** (194)	.54** (146)	.47** (235)	.56** (195)
/Task Feedback	.30** (145)	.21** (235)	.30** (194)	.30** (146)	.19** (235)	.25** (195)
/Required Skills	.76** (145)	.63** (235)	.79** (195)	.84** (146)	.85** (235)	.83** (195)
Autonomy						
/Task Completeness	.14 (143)	.32** (235)	.47** (194)	.62** (146)	.52** (235)	.61** (195)
/Task Feedback	.21** (144)	.29** (235)	.32** (194)	.35** (146)	.16* (235)	.31** (195)
/Required Skills	.45** (144)	.47** (235)	.52** (195)	.82** (146)	.80** (235)	.79** (195)

Table 13 (continued)

Multimethod Phase I Correlations for Five Job Dimensions, by Education
(N's appear in parentheses)

Dimension Pair	Interview			Observation		
	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)
Task Completeness						
/Task Feedback	.26** (143)	.21** (235)	.36** (194)	.60** (146)	.41** (235)	.57** (195)
/Required Skills	.09 (143)	.25** (235)	.32** (194)	.47** (146)	.42** (235)	.62** (195)
Task Feedback						
/Required Skills	.27** (145)	.20** (235)	.33** (194)	.38** (146)	.23** (235)	.42** (195)

*$p < 0.05$
**$p < 0.01$

Table 14

Multimethod Phase II Correlations for Five Job Dimensions, by Rated Intelligence
(N's appear in parentheses)

Dimension Pair	Interview		Observation	
	Low Intelligence	High Intelligence	Low Intelligence	High Intelligence
Variety				
/Autonomy	.29* (51)	.37** (59)	.71** (55)	.63** (57)
/Task Completeness	.03 (50)	.11 (59)	.58** (55)	.08 (57)
/Task Feedback	.02 (51)	.25 (58)	.27* (54)	-.08 (58)
/Required Skills	.30* (51)	.30* (59)	.51** (55)	.55** (58)
Autonomy				
/Task Completeness	.29* (51)	.22 (59)	.33* (55)	.01 (57)
/Task Feedback	.27 (52)	.39** (58)	.21 (54)	-.25 (57)
/Required Skills	.37** (52)	.40** (59)	.65** (55)	.70** (57)
Task Completeness				
/Task Feedback	.39** (51)	.47** (58)	.34** (54)	.39** (57)
/Required Skills	.31* (51)	-.18 (59)	.34** (55)	-.02 (57)
Task Feedback				
/Required Skills	.43** (52)	.10 (58)	.14 (54)	-.13 (58)

$*p < 0.05$
$**p < 0.01$

Table 15

Multimethod Phase II Correlations for Five Job Dimensions, by Education
(N's appear in parentheses)

Dimension Pair	Interview			Observation		
	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)
Variety						
/Autonomy	.21 (31)	.31* (50)	.62** (58)	.73** (34)	.74** (50)	.48** (58)
/Task Completeness	-.08 (31)	.11 (50)	.10 (57)	.59** (34)	.49** (50)	.11 (58)
/Task Feedback	.10 (31)	.14 (48)	.24 (58)	.37* (33)	.18 (50)	.20 (58)
/Required Skills	.23 (31)	.56** (50)	.30* (58)	.61** (34)	.54** (51)	.44** (58)
Autonomy						
/Task Completeness	.03 (32)	.36** (51)	.31* (57)	.44** (34)	.10 (50)	.08 (58)
/Task Feedback	.44** (32)	.43** (49)	.34** (58)	.29 (33)	.24 (49)	-.14 (58)
/ Required Skills	.31 (32)	.29* (51)	.47** (58)	.64** (34)	.75** (50)	.49** (58)

Table 15 (continued)

Multimethod Phase II Correlations for Five Job Dimensions, by Education
(N's appear in parentheses)

Dimension Pair	Interview			Observation		
	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)	Less Than High School Diploma (Low)	High School Diploma (Medium)	More Than High School Diploma (High)
Task Completeness						
/Task Feedback	.38* (32)	.43** (49)	.59** (57)	.53** (33)	.41** (49)	.23 (58)
/Required Skills	-.08 (32)	.03 (51)	.26* (57)	.66** (34)	.10 (50)	.05 (58)
Task Feedback						
/Required Skills	.32 (32)	.15 (49)	.28* (58)	.22 (33)	.14 (50)	.09 (58)

$*p < 0.05$
$**p < 0.01$

The correlations from Phase II shown in Tables 14 and 15 are less consistently ordered in the direction opposite to that predicted by hypothesis three. For intelligence, six of the correlational pairs differ in the unexpected direction, three in the expected direction, and the other pair of correlations (variety/required skills) are identical (Table 14). For the high and low levels of education, eight of the ten comparisons show higher correlations among the more highly educated workers, with only two of the correlational pairs running in the opposite (i.e., expected) direction (Table 15). These results are inconsistent with hypothesis three.

The correlations for observers in Tables 12-15 follow the pattern opposite to that found for respondents. Thus, objective task characteristics do not explain the correlational differences among respondents at various levels of cognitive sophistication. Specifically, the correlations among the observer-rated dimensions tend to be higher on the jobs of the less sophisticated respondents than on the jobs of the more sophisticated respondents, the trend being stronger for education (Tables 13 & 15) than for rated intelligence (Tables 12 & 14), and much stronger for Phase II (Tables 14 & 15) than for Phase I (Tables 12 & 13).

Discussion

The data in the present study support two of the three hypotheses concerning the impact of the cognitive sophistication of workers on their responses to an occupational survey. Hypothesis one predicts higher correlations between descriptive and importance ratings for various job facets among less sophisticated respondents, on account of their limited ability to abstract to hypothetical situations (e.g., an ideal job). For controls on both intelligence and education, the correlational differences are broadly in accord with the hypothesis in both phases for the overall measures and for the specific job facets. The predominantly favorable results for hypothesis one correspond to those reported by Moch et al. (1976) in which less sophisticated respondents evidenced a lower capacity to distinguish between actual and desired influence in two of three work domains.

In line with hypothesis two, the relationship between retrospective estimates of change and measured change is stronger overall among more sophisticated respondents, presumably because of their greater capacity to recall past events and changes. Hypothesis two receives only moderate support in the larger sample, which includes those who did and also those who did not change jobs. With the sample confined to workers who changed jobs, the data provide somewhat more convincing evidence in support of hypothesis two. As regards pertinent literature, the positive findings for hypothesis two parallel those reported in the study on recall of hospital episodes (U.S. Department of Health, Education, and Welfare, 1965).

The third hypothesis proposes that, because of their limited ability to discriminate, less cognitively sophisticated workers should generate higher intercorrelations among measures of different job facets. Contrary to hypothesis three, the intercorrelations among measures of facets are higher for workers classified as cognitively sophisticated than for those thought to be less sophisticated. This reversed correlational pattern proves stronger for Phase I than for Phase II.

The data on observers, who are presumed to provide a more objective estimate of the intercorrelations among job facets, conform to the correlational pattern predicted for workers, that is, higher intercorrelations among observers watching the jobs of less sophisticated workers. The correlational pattern obtained for observers emerges much more definitively in Phase II than in Phase I. The data on hypothesis three indicate, therefore, that objective task characteristics (as reported by observers) do not explain away the unexpected correlational pattern obtained for respondents.

Nonetheless, the particularly strong tendency in Phase II for relatively higher intercorrelations among the observers' ratings of the jobs of the less sophisticated workers may explain why the respondents' data are not ordered as consistently against hypothesis three in Phase II as they are in Phase I. As noted, observers provide relatively objective and workers relatively subjective ratings of the task dimensions of jobs.[2] The relatively higher interrelationships among dimensions of the jobs of the less sophisticated workers in Phase II may have limited the extent to which these workers could differentiate realistically among the dimensions.

---

[2] If, as suggested, the observers provide relatively objective ratings, the difference between the patterns of intercorrelations of observers' ratings at Phases I and II is puzzling. This may be understood as resulting from the particular samples used in Phases I and II. In Phase I, 579 jobs were observed, whereas in Phase II only 147 jobs were observed. Since these 147 jobs are not a random sample of the original 579 (and in fact include some jobs not observed at Phase I), there is no reason to expect that the pattern of intercorrelations would be the same for the two samples. When the observer data of the two phases are reanalyzed within a sample of jobs restricted to those jobs observed in both Phases I and II where the job occupant had not changed (N = 76), quite similar patterns of intercorrelations in the two phases are found. As in the larger sample, of course, the average correlation across jobs is smaller at Phase II, as a result of the training effort at Phase II to reduce halo effects among observers.

Placed in the context of existing literature, negative findings in the present study regarding hypothesis three square with the results obtained in the most comparable previous study (Moch et al., 1976). The latter study noted a trend toward higher correlations among sophisticated than unsophisticated respondents as regards their ratings of their influence in different domains of work. Studies that have successfully demonstrated greater halo effects among less sophisticated respondents (Gruenfeld & Arbuthnot, 1969; Schneier, 1977) have used measures of sophistication such as field-dependence and cognitive complexity that overlap only slightly with the intelligence dimensions studied here (Sechrest & Jackson, 1961; Vannoy, 1965; Vernon, 1972).

Although no obvious explanation emerges to account for the negative findings on hypothesis three, one possible interpretation may be cited. To a certain extent, the task of rating a job on five intrinsically rewarding job dimensions may generate pressure toward cognitive consistency. More intellectually sophisticated respondents may feel obliged to describe their jobs as favorable on one characteristic if they have already rated it high on another. Thus, they may display a more articulated "implicit organizational theory" than do their less sophisticated counterparts (cf. Eden and Leviatan, 1975, for a demonstration of the effect of an "implicit theory" of leadership). Less sophisticated respondents, conversely, may exhibit less awareness of, or concern about, the discrepancies among their ratings of various positive job attributes.

For all three hypotheses, highly similar results were obtained in the study for the two measures of cognitive sophistication: rated intelligence and level of education. This is not surprising, as the two measures intercorrelate almost as highly as the reliability of the ratings of intelligence by interviewers permits. The sources of the discrepancies between the results obtained for intelligence and for education are difficult to pinpoint, since the analyses differ not only in terms of the control variable used (and its reliability), but also with respect to sample size, number of categories, and extremity of groups compared.

In conclusion, the results obtained in this study have implications for future occupational and perhaps other surveys. Cognitively sophisticated and unsophisticated respondents appear to differ in at least two cognitive operations, abstraction and recall. Sometimes these cognitive differences lead to higher correlations among the more sophisticated respondents (hypothesis two), as was true of Converse's (1964) political research; at other times the correlations are higher among the less intellectually sophisticated (hypothesis one). Thus, when survey researchers ask questions that require such cognitive operations as abstraction or recall, they would do well to consider whether respondent sophistication may affect the answers they receive, as well as the relationships among these answers. If so, they should perhaps consider conducting their statistical analyses separately for the subsamples of cognitively sophisticated and unsophisticated respondents.

References

Barnowe, J. T., Mangione, T. W., & Quinn, R. P. The relative importance of job facets as indicated by an empirically derived model of job satisfaction. Ann Arbor, Mi.: Survey Research Center, 1972 (mimeo).

Bieri, J. Complexity-simplicity as a personality variable in cognitive and preferential behavior. In D. W. Fiske & S. Maddi (Eds.), *Functions of varied experience*. Homewood, Ill.: Dorsey, 1961.

Bieri, J., Atkins, A. L., Briar, S., Leaman, R. L., Miller, H., & Tripodi, T. *Clinical and social judgement*. New York: Wiley, 1966.

Bishop, G. F. The effect of education on ideological consistency. *Public Opinion Quarterly*, 1976, *40*, 337-348.

Converse, P. E. The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent*. New York: The Free Press, 1964.

Eden, D., & Leviatan, U. Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *Journal of Applied Psychology*, 1975, *60*, 736-741.

Gruenfeld, L., & Arbuthnot, J. Field independence as a conceptual framework for prediction of variability in ratings of others. *Perceptual and Motor Skills*, 1969, *28*, 31-44.

Hackman, J. R., & Lawler, E. E. III. Employee reactions to job characteristics. *Journal of Applied Psychology*, 1971, *55*, 259-286.

Jenkins, G. D., Jr., Nadler, D. A., Lawler, E. E. III, & Cammann, C. Standardized observations: An approach to measuring the nature of jobs. Journal of Applied Psychology, 1975, 60, 171-181. Also appears as Chapter 4.

Lord, F. M. Elementary models for measuring change. In C. W. Harris (Ed.), Problems in measuring change. Madison: University of Wisconsin, 1963.

Moch, M., Cammann, C., & Cooke, R. A. A subsystem approach to the measurement of influence in organizations. Paper presented at the Seventy-first Annual Meeting of the Americal Sociological Association, New York, August, 1976.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Quinn, R. P., & Cobb, W. L., Jr. What workers want: Factor analyses of importance ratings of job facets. Ann Arbor, Mi.: Survey Research Center, 1971 (mimeo).

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey Ann Arbor, Mi.: Survey Research Center, 1974.

Schneier, C. A. Operational utility and psychometric characteristics of behavioral expectation scales: A cognitive reinterpretation. Journal of Applied Psychology, 1977, 62, 541-548.

Schuman, H., & Presser, S. Question wording as an independent variable in survey analysis. Sociological methods and research, 1977, 6, 151-176.

Sechrest, L., & Jackson, D. N. Social intelligence and accuracy of interpersonal predictions. Journal of Personality, 1961, 29, 167-182.

United States Department of Health, Education, and Welfare. Reporting of hospitalization in the health interview survey. (Vital and Health Statistics, Series 2, No. 6, 1965).

Vannoy, J. S. Generality of cognitive complexity-simplicity as a personality construct. *Journal of Personality and Social Psychology*, 1965, *2*, 385-396.

Vernon, P. E. The distinctiveness of field independence. *Journal of Personality*, 1972, *40*, 366-391.

Chapter 12

ROLE AMBIGUITY AND ROLE STRAIN: A SEARCH FOR INTERVENING VARIABLES

by

Terry A. Beehr

ABSTRACT

This study sought to clarify the question why role ambiguity is related to aversive responses by role occupants--a finding often reported in previous research. The model tested was a simple sequential chain of causation in which role ambiguity is hypothesized to cause condition X (an intervening variable) which in turn causes an aversive response. This chain model was tested with two possible intervening variables suggested by prior research, and five outcome, or criterion, variables representing role strain.

The results provide only weak and partial support for such a model in the case of one hypothesized intervening variable (role overload) and no support at all in the case of a second intervening variable (skill underutilization). The data also suggest that there may be a direct causal relationship between role ambiguity and role strain, without intervention of other variables.

It is proposed that the chain-of-causation model as specified or implied in prior research is inappropriate to the problem and that more complex models must be invoked. The nature of these alternative models is suggested. A simple linear additive model of causation appeared to fit the data employed in this analysis.

Chapter 12

ROLE AMBIGUITY AND ROLE STRAIN: A SEARCH FOR INTERVENING VARIABLES

There is much evidence that ambiguous roles in work organizations are associated with aversive psychological and physiological states of the role occupant (Beehr, 1974; Caplan & Jones, 1974; French & Caplan, 1973; Hamner & Tosi, 1974; House & Rizzo, 1972; Kahn et al., 1964; Lyons, 1971). It has further been shown that for employees with a low tolerance for ambiguity, this association is particularly strong (Beehr, 1974; Lyons, 1971; Kahn et al., 1964).

The question of why ambiguity in organizational roles is stressful has been investigated in very few studies. The usual assumption is that the ambiguity itself is tension-producing (Sinha, 1950). Indeed, it has been argued that the tendency of some people to seek reduction of ambiguity in their environment is evidence for such aversion to ambiguity (Adorno et al., 1950; Frenkel-Brunswick,1949; Millon, 1957; Smock, 1955). Ambiguity reduction may take a variety of forms, such as stereotyping, resisting change, and speed in seeing clear shapes within ambiguous visual stimuli. Withdrawal from ambiguous situations is another response indicating aversion to ambiguity. Quinn, Fine, and Levitin (1970) showed that ambiguity regarding what a worker's daily work routine would be was related to turnover among new blue collar employees at an automobile plant, and

Lyons (1971) found that role ambiguity was related to subsequent voluntary turnover among registered nurses at a large hospital.

Aside from direct aversive effects of role ambiguity, it has been proposed that other variables may intervene in causal chains between role ambiguity and role strain in work roles. French and Caplan (1973) found that role ambiguity was related to the underutilization of the skills and abilities of male administrators, engineers, and scientists at Goddard Space Flight Center. Obviously, if it is not clear how an employee should use his skills, he will not be able to use them fully. Because underutilization of skills was related to role strain in that study (also in Margolis, Kroes & Quinn, 1974), French and Caplan suggested a model in which part of the aversive effect of role ambiguity is due to its effect on the underutilization of skills. According to the model, role ambuiguity leads to underutilization of skills which leads to role strain.

Kahn and Quinn (1970) hypothesized that role ambiguity is related to individual role strain through another intervening variable, role overload. They argue that role overload "is one potential by-product of role ambiguity, particularly where the ambiguous expectations are evaluational: the role occupant must constantly keep doing 'more' to be certain that he has at least done 'enough'". Role overload, having more work to do than can be done in the time available, can result in such a situation, and role overload has been shown related to role strain among employees and students (French & Caplan, 1973; Caplan & Jones, 1974).

This paper reports a test of three hypotheses. The first hypothesis is that role ambiguity is related to role strain through its influence on role overload; the second hypothesis is that role ambiguity is related to strain through its influence on underutilization of skills; and the third hypothesis is that role ambiguity is directly related to role strain, i.e., it has some effect on role strain in addition to any effect it might have through role overload and underutilization of skills.

Method

Sample

The sample comprised 649 people employed full-time (35 hours or more per week) by five Midwestern work organizations: a printing company, a small research and development company, two automotive supply companies, and the four services departments of a hospital. All supervisors were in the sample, and non-supervisors were systematically sampled at rates varying from 25 percent to 100 percent. The response rate was 72.9 percent.

Measures

Data were collected during a 90 minute structured interview in the respondent's home.

Role ambiguity was measured by a four-item index ($r_{kk} = .69$).* Examples of items are: "I can predict what others will expect of me tomorrow", and "I am clear what others expect of me on my job". The items were rated from "very true" to "not at all true" on a four-point scale.

Five role strains were measured. Job dissatisfaction was measured by a four-item subset ($r_{kk} = .80$) of Quinn and Shepard's (1974) global job satisfaction index. A typical item was "All in all, how satisfied would you say you are with your job?" rated from "very satisfied" to "not at all satisfied" on a four-point scale. The life dissatisfaction measure (r_{kk} = .79) consisted of a nine-item subset of Quinn and Shepard's (1974) life satisfaction index. There were seven seven-point semantic differential items, e.g., "discouraging-hopeful", and two fixed-alternative questions, e.g., "Taking all things together, how would you say things are these days? Would you say you're very happy, pretty happy, or not too happy these days?" Low self-esteem was measured by a three-item subset (r_{kk} = .68) of Quinn and Shepard's (1974) seven-point semantic differential scale. An example is "successful-unsuccessful". The depressed mood index (r_{kk} = .71) was a ten-item index taken from Quinn and Shepard (1974). A sample item is, "I feel down-hearted and blue", rated on a four-point scale from "often" to "never". Fatigue (r_{kk} = .71) was measured by three items, e.g., "In the last year how often did you feel completely worn out at the end of the day?" rated on a four-point scale from "often to "never" and a fourth item asking the respondent to indicate which step on a seven-step ladder indicated "how much pep and energy (you've) had lately". The top step represented always being full of pep and energy, and the bottom step indicated never having any pep or energy.

---

*r_{kk} = Spearman-Brown index reliability estimate based upon median item intercorrelations.

Role overload was measured by two items (r_{kk} = .47), e.g., "I am given enough time to do what others expect of me", rated on a four-point scale from "very true" to "not at all true".

Underutilization of skills was measured by a three-item index. The first item, the level of education (on an eight point scale from "none" to "graduate or professional training") required to do the employee's job, was subtracted from the second item, the level of education of the employee (on the same eight-point scale). This difference was added to the third item, "Through your previous experience and training, do you have some skills that you would like to be using in your work but can't use on your present job?" coded seven for "yes" and zero for "no".

The intercorrelations among these measures appear in Table 1.

Analyses

The first two hypotheses were tested by finding the significance of the difference between the following correlations: the zero-order correlation between role ambiguity and the corresponding role strain with the effect of the proposed intervening variable held constant. If the third hypothesis was correct this partial correlation would be significant.

Results

The first column of Table 1 shows that role ambiguity was related to each of the five role strains and also related ($r = .45$, $p < .01$) to role overload. The second column of Table 1 shows that role overload was related to each role strain. Table 2 shows that the differences between the zero-order correlations (ambiguity-strains) and the partial correlations (effects of overload removed) were large enough to suggest weak support for the first hypothesis for three of the five role strains: job dissatisfaction ($p < .05$), low self-esteem ($p < .10$), and fatigue ($p < .10$).

While underutilization of skills was related to four of the five roles strains (see Table 1), it was not significantly related to role ambiguity ($r = .07$), and therefore, it could not act as an intervening variable

Table 1

Intercorrelations among All Variables
(N = 649)

	1	2	3	4	5	6	7	8
1. Role Ambiguity	(.69)[a]							
2. Role Overload	.45	(.47)						
3. Skill Underutilization	.07[b]	.05[b]	-					
4. Job Dissatisfaction	.22	.27	.27	(.80)				
5. Life Dissatisfaction	.09	.10	.06	.27	(.79)			
6. Low Self-esteem	.19	.18	.14	.25	.28	(.68)		
7. Depressed Mood	.16	.16	.14	.43	.39	.31	(.71)	
8. Fatigue	.12	.18	.12	.22	.36	.19	.36	(.71)

[a] Figures in parentheses are reliability estimates (Spearman-Brown) based upon median item intercorrelations.

[b] Not significant; all other correlations are significant, $p < .01$.

Table 2

The Relationship of Role Ambiguity to Role Strain through Role Overload

Role Strain	r_1 Correlation between role ambiguity and role strain (N = 614)	r_2 Correlation between role ambiguity and role strain with role overload held constant (N = 628)	Significance of difference between r_1 & r_2
Job dissatisfaction	.22**	.11**	$p < .05$
Life dissatisfaction	.09**	.05	n.s.
Low self-esteem	.19**	.12**	$p < .10$
Depressed mood	.16**	.10**	n.s.
Fatigue	.12**	.04	$p < .10$

*$p < .01$

between role ambiguity and role strain in a causal chain. Thus, the second hypothesis was not supported.

Since the second hypothesis was not supported by the data, the testing of the third hypothesis required that only role overload be held constant in computing the partial correlations between role ambiguity and role strain. The second column of Table 2 shows that these partial correlations were greater than zero for three of the five role strains: job dissatisfaction ($p < .01$), low self-esteem ($p < .01$), and depression ($p < .01$). Thus, the third hypothesis was supported for these role strains.

Discussion

This study sought to clarify the question, why role ambiguity is related to aversive consequences for role occupants. The model tested was a simple sequential chain of causation in which role ambiguity resulted in X (an intervening variable) which in turn was experienced as being aversive by each of several strain criteria. This chain model, and two proposed intervening variables, were examined with reference to five strain criteria.

The results provide only weak and partial support for such a model in the case of one intervening variable, role overload, and no support at all in the case of the other, i.e., underutilization of skills. The results also suggest that there may be a direct causal relationship between role ambiguity and role strain, although the evidence is slight and omitted consideration of other possible intervening variables.

It is clear that the simple chain of causation model, implied or asserted in previous research, is inappropriate to the problem, and that a more complex model must be invoked. Such a more complex model may well provide for a set of causal factors operating in an additive, substitutive, or interactive way to generate the role strain outcomes. In addition, there is the possibility that the role strains, here defined as outcomes, may also operate as "causes" as in the case, for example, of fatigue itself causing or enhancing role overload. In any case, the exploration of such more complex models appeared inappropriate to the available data which are limited to concurrent measures and have the defect of possible error from correlated response bias. An application of multiple regression analysis to the present data, however, suggests that a simple linear additive model is sufficient to account for nearly all of the criterion variance available for prediction with the set of hypothesized causal variables of role ambiguity, role overload, and skill underutilization. The relevant data are summarized in Table 3.

Table 3

Multiple and Partial Correlations between Three Role Strain Predictors and Five Role Strain Criteria

Role Strain Criteria	Partial Correlations			Multiple Correlation
	Role Ambiguity	Role Overload	Skill Under-Utilization	
Job Dissatisfaction	.11**	.19**	.30**	.41**
Life Dissatisfaction	.05	.07	.06	.14**
Low Self-Esteem	.12**	.10**	.13**	.25**
Depressed Mood	.10**	.09**	.13**	.23**
Fatigue	.04	.14**	.13**	.22**

**$p < .01$

References

Adorno, T., Frenkel-Brunswik, E., Levinson, D., & Sanford, R. N. The *authoritarian personality*. New York: Harper, 1950.

Beehr, T. A. Role ambiguity as a role stress: Some moderating and intervening variables. Doctoral dissertation, The University of Michigan, 1974.

Caplan, R. D., & Jones, K. W. Technological dependency, role stress, and strain. Paper presented at the annual meeting of The American Psychological Association Convention. New Orleans, September, 1974.

French, J. R. P., Jr., & Caplan, R. D. Organizational stress and individual strain. In A. J. Marrow (Ed.), *The failure of success*. New York: AMACOM, 1973

Frenkel-Brunswik, E. Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality*, 1949, *18*, 108-143.

Hamner, W. C., & Tosi, H. L. Relationship of role conflict and role ambiguity to job involvement measures. *Journal of Applied Psychology*, 1974 *59*, 497-499.

House, R. J., & Rizzo, J. R. Role conflict and ambiguity as critical variables in a model of organizational behavior. *Organizational Behavior and Human Performance*, 1972, *7*, 467-505.

Kahn, R. L., & Quinn, R. P. Role stress: A framework for analysis. In A. McLean (Ed.), *Occupational mental health*. New York: Rand McNally, 1970.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. *Organizational stress: Studies in role conflict and ambiguity*. New York: Wiley, 1964.

Lyons, T. F. Role clarity, need for clarity, satisfaction, tension, and withdrawal. Organizational Behavior and Human Performance, 1971, 6, 99-110.

Margolis, B. L., Kroes, W. H., & Quinn, R. P. Job stress: An unlisted hazard. Journal of Occupational Medicine, 1974, 10, 659-661.

Millon, T. Authoritarianism, intolerance of ambiguity, and rigidity under ego- and task-involving conditions. Journal of Abnormal and Social Psychology, 1957, 55, 29-33.

Quinn, R. P., Fine, B. D., & Levitin, T. Turnover and training: A social-psychological study of disadvantaged workers. Survey Research Center, Ann Arbor: Report to the Manpower Administration, September 1970.

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey: Descriptive statistics with comparison data from the 1969-70 Survey of Working Conditions. Ann Arbor: Survey Research Center, 1974.

Sinha, A. Experimental induction of anxiety by conditions of uncertainty. Doctoral dissertation, The University of Michigan, 1950.

Smock, C. D. The influence of psychological stress on the "intolerance of ambiguity". Journal of Abnormal and Social Psychology, 1955, 50, 177-182.

Chapter 13

PERCEIVED SITUATIONAL MODERATORS OF THE RELATIONSHIP BETWEEN SUBJECTIVE ROLE AMBIGUITY AND ROLE STRAIN

by

Terry A. Beehr

ABSTRACT

Previous research has focused on personality moderators of the relationship between organizational stress and individual strain. This study was a search for situational moderators of the relationship between one organizational stress, role ambiguity, and four psychological strains: job dissatisfaction, life dissatisfaction, low self-esteem, and depressed mood. Three situational characteristics were hypothesized to moderate the relationship by reducing its strength: group cohesiveness, supervisor support, and autonomy. Group cohesiveness moderated the relationship between role ambiguity and two of the role strains, but the direction of its moderating influence was inconsistent. An explanation was offered for this result. Supervisor support showed a nonsignificant tendency to reduce the strength of the relationship between role ambiguity and role strain. Autonomy tended to moderate the relationship in the expected direction significantly and strongly.

Chapter 13*

PERCEIVED SITUATIONAL MODERATORS OF THE RELATIONSHIP BETWEEN SUBJECTIVE ROLE AMBIGUITY AND ROLE STRAIN**

Role stress is any aspect of role expectations that has adverse consequences (role strain) for individual role incumbents (Kahn & Quinn, 1970). Most research on role stress has focused on samples representative of limited populations, for example, a single role or a small number of organizational roles (French & Caplan, 1973; Lyons, 1971; Paul, 1974; Rizzo, House, & Lirtzman, 1970; Caplan & Jones, 1974).

In addition to the use of narrow samples, most previous research has been limited to the investigation of the direct, overall relationship between role stress and role strain. A few researchers have tested personality characteristics (Johnson & Stinson, 1975; Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964; Lyons, 1971; Caplan & Jones, 1974) as

---

* This paper first appeared in the Journal of Applied Psychology, 1976, 61(1), 35-40.

** This study is based on the author's doctoral dissertation, "Role Ambiguity as a Role Stress: Some Moderating and Intervening Variables," which was completed at The University of Michigan.

The author wishes to thank Edward Bordin, Robert Cooke, Jack Drexler, Larry Michaelsen, Robert Quinn, and Stanley Seashore for their help during various stages of work on this article.

variables in a role stress model. There has been no serious attempt, however, to test situational characteristics as moderators of the relationship between role stress and strain.

In this study, three situational characteristics were hypothesized as moderators of the relationship between role stress and role strain. Group cohesiveness and supervisor support were proposed as moderators because of the argument by Kahn and Quinn (1970) that psychological support in the presence of stress may reduce role strain. (Presumably, cohesive groups are more supportive than noncohesive groups.) Autonomy also was hypothesized as a moderator. If the role expectations that others send to the focal person are stressful, people with job autonomy may define their own role expectations, resulting in reduced role strain.

This study added to the knowledge from previous studies of role stress in two ways. First, the variety of roles and of organizations was greater, so that the results would be representative of phenomena that are widely dispersed in the world of work. Second, situational rather than personality moderators were examined. Role ambiguity, that is, uncertainty regarding what is expected on one's job, was chosen for analysis because of the long history of psychological research on ambiguity, and because of the Kahn et al. (1964) finding that role ambiguity is felt to a stressful extent by about one third of the American male wage and salary workers. Four psychological role strains were measured: job dissatisfaction, life dissatisfaction, low self-esteem, and depressed mood.

The specific hypotheses were that role ambiguity is more strongly related to the role strains (a) among people in noncohesive groups than among people in cohesive groups, (b) among people with nonsupportive supervisors than among people with supportive supervisors, and (c) among people in nonautonomous roles than among people in autonomous roles.

Method

Sample

The sample comprised 651 respondents employed 35 hours or more per week by five midwestern work organizations: a printing company (n = 173), a small research and development company (n = 21), two automotive supply companies (ns = 120 and 124), and four services departments of a hospital (n = 213). These organizations were chosen in part on the criterion that they contain a diverse set of jobs, and there is evidence that the occupational census codes of the respondents roughly match the codes of a national sample of workers (Beehr, 1974). All supervisors were included in the sample, and nonsupervisors were sampled systematically at rates varying from organization to organization (between 25% and 100%). The response rate was 72.9%.

Fifty-one percent of the respondents were male, 79.8% were white, 68.5% were married, 73.0% had finished at least high school, and the mean age was 35 years. Comparable information on the nonrespondents was not available.

Measures

Data were collected during a 90-minute structured interview in the respondent's home conducted during the winter of 1972-1973 by interviewers from the Survey Research Center.

The measures of all variables in the study were subjective, that is, they were perceptions of the subjects. It could be argued, however, that the perceptions of the subject are important in stress research, because strain may be largely an individual's reaction to his subjective environment. In addition, some of the environmental variables in this study have a partially subjective nature. For example, regardless of a supervisor's objective behavior, it cannot be said that a subordinate is given true psychological support unless the subordinate feels supported.

Since all of the data were gathered during a single interview, there is a possibility that relationships could be inflated by correlated method variance. Since the present study is seeking differences in the strengths of relationships, however, the results probably are conservative, that is, more significant differences might have been found if multiple sources of data had been used.

Role ambiguity (M = 1.92, SD = .64) was measured by a four-item index:

How true is this of your job?

> I can predict what others will expect of me tomorrow.
>
> I am clear on what others expect of me on my job.
>
> On my job, whatever situation arises, there are procedures for handling it.
>
> I get enough facts and information to work my best.

All items except the third (On my job . . .) were taken from the Quality of Employment Survey (Quinn & Shepard, 1974). The response categories were "very true," "somewhat true," "a little true," and "not at all true."

Four psychological strains were measured. Presumably, psychological strains are more accurately measured by interviews than physiological strains would be. The first, job dissatisfaction (M = 3.94, SD = .97), was measured by a four-item subset of Quinn and Shepard's (1974) global job satisfaction index:

> All in all, how satisfied would you say you are with your job--very satisfied, somewhat satisfied, not too satisfied, or not at all satisfied?
>
> In general, how well would you say that your job measures up to the sort of job you wanted when you took it? Would you say it is very much like, somewhat like, or not very much like the job you wanted when you took it?
>
> Knowing what you know now, if you had to decide all over again whether to take the job you now have, what would you decide? Would you decide without any hesitation to take the same job, would you decide definitely not to take the same job?
>
> If a good friend of yours told you that he/she was interested in working in a job like yours, what would you tell him/her? Would you strongly recommend this job, would you have doubts about recommending it, or would you strongly advise him/her against this sort of job?

The job dissatisfaction index had a possible range of 1-5.

The life dissatisfaction measure (M = 4.83, SD = 1.25) consisted of a nine-item subset of Quinn and Shepard's (1974) life satisfaction index.

Seven of these items were answered on a semantic differential scale. The seven pairs of anchors were "interesting-boring," "enjoyable-miserable," "friendly-lonely," "full-empty," "hopeful-discouraging," "rewarding-disappointing," and "brings out the best in me-doesn't give me much of a chance" in response to the question, "How do you feel about your present life in general?" The fixed-alternative questions were:

> In general, how satisfying do you find the ways you're spending your life these days? Would you call it completely satisfying, pretty satisfying, or not very satisfying?
>
> Taking all things together, how would you say things are these days? Would you say you're very happy, pretty happy, or not too happy these days?

The life dissatisfaction index had a possible range of 1-7.

Low self-esteem (M = .99, SD = 1.05) was measured by a three-item subset of Quinn and Shepard's (1974) 7-point semantic differential scale. The anchors were "successful-unsuccessful," "important-not important," and "doing my best-not doing my best" in response to the question, "How do you see yourself in your work?"

Depressed mood (M = 1.85, SD = .47) was measured by a 10-item index taken from Quinn and Shepard (1974).

How often do you feel this way at work?

I feel downhearted and blue.
I get tired for no reason.
I find myself restless and can't keep still.
I find it easy to do the things I used to do.
My mind is as clear as it used to be.
I feel hopeful about the future.

I find it easy to make decisions.

I am more irritable than usual.

I still enjoy the things I used to.

I feel that I am useful and needed.

The response categories were "often," "sometimes," "rarely," and "never."

Three situational characteristics, that is, group cohesiveness, supervisor support, and autonomy, were measured. The first, group cohesiveness (M = 3.65, SD = 1.19), was measured by five items:

Is there any group of people at [name of respondent's organization] that you think of as your coworkers--people whom you see just about everyday and with whom you have to work closely in order to do your job well? I know it's hard to talk about a whole group of people like this, but I'd like to get your general views about the people in this group.

How ready are they to defend each other from criticism by outsiders?

How well do you feel they help each other on the job?

How well do they get along together?

How well do they stick together?

The last three items were adapted from Seashore (1954). The response categories for the last four items were "great, couldn't be better," "pretty good, but some could be better," "not very good, but some pretty good," and "not very good." Respondents answering no to the first question were not asked the other four questions and were assigned a score of 1 on the group cohesiveness index. The response categories for the last four questions were coded 2 through 5, and respondents answering those questions received their mean score on those four items as their score on the group cohesiveness measure. Thus, a high score on the index meant that the respondent was part of a very cohesive group, and a low score indicated that he was part of no group at all or of a group with little cohesiveness.

Supervisor support (M = 3.12, SD = .81) was measured by three items:

How true is this of your supervisor?

My supervisor keeps informed about the way his/her people think and feel about things.

My supervisor stands up to outsiders for the people he/she supervises.

My supervisor takes a personal interest in those he/she supervises.

These items were taken from Hemphill and Coombs (1957), and the response categories were "very true," "somewhat true," "a little true," and "not at all true."

Autonomy (M = 2.97, SD = .76) was measured by four items:

How true is this of your job?

I have a lot of say over what happens on my job.

I have enough authority to do my best.

My job allows me to make a lot of decisions on my own.

I have enough freedom as to how I do my work.

These items were taken from Quinn and Shepard (1974), and the response categories were "very true," "somewhat true," "a little true," and "not at all true." Table 1 contains the correlations among the variables for the whole sample. Although all of the variables were obtained from the same interview, none of the correlations was strong enough to suggest that any two variables were measuring identical constructs.

Analyses

The sample was divided at the median on each situational variable, and Pearson product-moment correlations between role ambiguity and each role strain were computed within each half sample. Hypotheses were directional, that is, the relationship between role ambiguity and each role strain was predicted to be stronger in the less cohesive than in the more cohesive group; stronger in the low-supervisor-support than in the high-supervisor-support group; and stronger in the low-autonomy than in the high-autonomy group. Therefore, one-tailed tests of the significance of the difference between correlations were computed.

Results

The first column of Table 1 shows that role ambiguity is significantly related to the four role strains for the sample as a whole. Job dissatisfaction is the strain most strongly correlated with role ambiguity. Life dissatisfaction, as one would expect, has the weakest relationship to ambiguity in the work role. Obviously, there are many sources of life satisfaction other than those in the work role.

Table 2 shows that group cohesiveness moderates the relationship between role ambiguity and low self-esteem as predicted ($p < .05$). However, group cohesiveness moderates the relationship between role ambiguity and job dissatisfaction ($p < .10$) in the direction opposite from that predicted. A two-tailed test was used for this comparison because the direction had not been predicted correctly.

Although three of the four differences between correlations are in the predicted direction for the analysis of supervisor support as a moderator,

Table 1

Correlations of Role Ambiguity, Strains, and Situational Moderators

Item	1	2	3	4	5	6	7	8
Role stress								
1. Role ambiguity	(.69)							
Role strains								
2. Job dissatisfaction	.22**	(.80)						
3. Life dissatisfaction	.08*	.27**	(.79)					
4. Low self-esteem	.19**	.25**	.28**	(.68)				
5. Depressed mood	.19**	.43**	.39**	.31*	(.71)			
Situational moderators								
6. Group cohesiveness	-.12**	-.14**	-.02	-.03	-.05	(.78)		
7. Supervisor support	-.30**	-.37**	-.17**	-.23**	-.32**	.13**	(.78)	
8. Autonomy	-.26**	-.37**	-.21**	-.27**	-.32**	.13**	.36**	(.74)

Note. N = 587 and represents the number of respondents having complete data for all variables in the table. Numbers in parentheses are reliability estimates obtained by applying the Spearman-Brown prophecy formula to the median interitem correlation.

*p < .05.
**p < .01.

Table 2

Group Cohesiveness as a Moderator of the Relationship between Role Ambiguity and Psychological Role Strain

	Correlations with Role Ambiguity		
	r_1	r_2	Significance of difference between
Role strain	Low group cohesiveness	High group cohesiveness	r_1 and r_2
Job dissatisfaction	.16*	.29*	$p < .10$
Life dissatisfaction	.05	.15*	ns
Low self-esteem	.24*	.08	$p < .05$
Depressed mood	.19*	.10	ns

Note. For r_1, n = 386, for r_2, n = 246; ns represent the number of respondents having complete data for all variables in the table.

*$p < .01$.

none of the differences is statistically significant.

All differences between correlations of role ambiguity and role strain are in the predicted direction when the sample is divided on autonomy (see Table 3). The differences for job dissatisfaction and depression are significant at the .01 level and the difference for low self-esteem at the .10 level.

Discussion

This study suggested that even if there is a stress (role ambiguity), people in roles with certain situational characteristics, especially autonomy, do not suffer from it. This has important implications because situational characteristics can be modified by organizations.

There was only suggestive evidence that people with supportive supervisors might not feel some role strains even if their roles are ambiguous. Even if psychological support has the expected beneficial effect, perhaps the supervisor is not the best source of that support. Previous laboratory work (Schachter, 1959) has shown that people expecting an electrical shock prefer to spend their waiting time with others like themselves, especially with others expecting shocks. This indicates that peers, especially those in similarly stressful roles, may be the most beneficial source of psychological support.

Group cohesiveness has an effect on the extent to which role ambiguity is related to strain, but the results are contradictory. For persons in cohesive groups, the relationship between role ambiguity and low self-esteem is weaker than it is for those in noncohesive groups. Job dissatisfaction, however, is related more strongly to role ambiguity in cohesive than in noncohesive groups.

Table 3

Autonomy as a Moderator of the Relationship between Role Ambiguity and Psychological Strain

	Correlations with Role Ambiguity		
	r_1	r_2	Significance of difference between r_1 and r_2
Role Strain	Low autonomy	High autonomy	
Job dissatisfaction	.25*	.05	$p < .01$
Life dissatisfaction	.10	.03	ns
Low self-esteem	.19*	.09	$p < .10$
Depressed mood	.23*	-.02	$p < .01$

Note. For r_1, n = 340, for r_2, n = 296; ns represent the number of respondents having complete data for all variables in the table.

*$p < 0.01$

A partial explanation of these results stems from the assumption that people in cohesive groups are more likely to communicate with coworkers than are people in noncohesive groups. People perceiving ambiguous roles can blame the ambiguity either on their roles or on themselves. Discussing the work situation with peers may result in social support for external placement of the blame, that is, on the role rather than on the self. Thus, while people in cohesive groups still are dissatisfied with ambiguous roles, they are less likely to internalize the blame in the form of lowered self-esteem. People in noncohesive groups are not as likely to have such discussions and, therefore, are more likely to place the blame for the ambiguity on themselves. The fact that both dissatisfaction measures, that is, job dissatisfaction and life dissatisfaction, tend to be correlated more strongly with role ambiguity in the high-cohesive group while low self-esteem and depressed mood tend to be correlated more strongly with role ambiguity in the low-cohesive group adds confidence to this interpretation.

Autonomy is the strongest and most consistent moderator of the relationship between role ambiguity and role strain. Organizations that wish to reduce the role strain associated with role ambiguity should increase the autonomy in their members' roles. Future research should attempt to determine whether job characteristics such as autonomy reduce the relationship between a large number of role stresses and role strain. If they do, redesigning jobs would seem even more likely to reduce the strain due to role stress.

References

Beehr, T. A. Role ambiguity as a role stress: Some moderating and intervening variables. Unpublished doctoral dissertation, University of Michigan, 1974.

Caplan, R. D., & Jones, K. W. Technological dependency, role stress, and strain. Paper presented at the meeting of the American Psychological Association, New Orleans, September 1974.

French, J. R. P., & Caplan, R. D. Organizational stress and individual strain. In A. J. Marrow (Ed.), The failure of success. New York: AMACOM, 1973.

Hemphill, J. K., & Coons, A. E. Development of the leader behavior description questionnaire. In R. M. Stogdill & A. E. Coons (Eds.), Leader behavior: Its description and measurement (Research Monograph No. 88). Columbus: Bureau of Business Research, Ohio State University, 1957.

Johnson, T. W., & Stinson, J. E. Role ambiguity, role conflict, and satisfaction: Moderating effects of individual differences. Journal of Applied Psychology, 1975, 60, 329-333.

Kahn, R. L., & Quinn, R. P. Role stress: A framework for analysis. In A. McLean (Ed.), Occupational mental health. New York: Rand McNally, 1970.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. Organizational stress: Studies in role conflict and ambiguity. New York: Wiley, 1964.

Lyons, T. F. Role clarity, need for clarity, satisfaction, tension, and withdrawal. Organizational Behavior and Human Performance, 1971, 6, 99-110.

Paul, R. J. Role clarity as a correlate of satisfaction, job related strain and propensity to leave--male vs. female. The Journal of Management Studies, 1974, 11, 233-245.

Quinn, R. P., & Shepard L. J. The 1972-1973 quality of employment survey: Descriptive statistics with comparison data from the 1969-1970 survey of working condtions. Ann Arbor, Michigan: The Institute for Social Research, 1974.

Rizzo, J. R., House, R. J., & Lirtzman, S. K. Role conflict and ambiguity in complex organizations. Administrative Science Quarterly, 1970, 15, 150-163.

Schachter, S. The psychology of affiliation. Stanford, Calif.: Stanford University Press, 1959.

Seashore, S. E. Group cohesiveness in the industrial work group. Ann Arbor, Michgian: The University of Michigan Press, 1954.

Chapter 14

WHAT MAKES JOBS MONOTONOUS AND BORING

by

Robert P. Quinn

ABSTRACT

Programs for job enlargement and enrichment commonly rest upon the assumption that increased variety of work activities will reduce the monotony of jobs and the resulting degree of experienced boredom. The available evidence is fragmentary and mixed, with suggestions that personal attributes of the worker moderate the presumed associations.

This inquiry attempted to clarify the matter by asking three questions: (1) What is the degree of contribution of objective task characteristics to the perception of monotony and experience of boredom? (2) What aspects of the temporal configuration of tasks are most influential in relation to monotony and boredom?, and (3) Are these relationships moderated by the worker's intelligence and/or his level of off-the-job activities?

The task characteristics examined were five in number, each representing some aspect of task variety; these task characteristics were assessed by trained observers without input from the job occupants. Monotony was assessed by observer ratings and also by report of the perceptions of job occupants. Boredom was regarded as a subjective experience of the job occupants and measured by self report.

All five indicators of task variety were found to be associated with monotony, and three were found associated with workers' reports of boredom. Intelligence had no significant effect upon the relationship between variety and monotony or boredom, but, along with level of off-the-job activities, did moderate the association between monotony and boredom.

These results are interpreted in the context of theories concerning the nature of boredom and implications for the design of boredom-reducing jobs.

Chapter 14

WHAT MAKES JOBS MONOTONOUS AND BORING?

Experiments with job enrichment and job enlargement as ways of increasing workers' performance or job satisfaction lead to two conclusions. First, published reports of such experiments indicate that they have not been unqualified successes (Hulin & Blood, 1969; Lawler, 1969; Quinn, Staines, & McCullough, 1974). Even this judgment does not take into account a large number of experiments that, because they were unsuccessful, went unreported. Second, even the successes have often been restricted to particular working populations. Hulin and Blood observe, for example, that

> . . . the argument for larger jobs as a means of motivating workers, decreasing boredom and dissatisfaction, and increasing attendance and productivity is valid only when applied to certain segments of the work force--white-collar and supervisory workers and nonalienated blue-collar workers (1969, p. 388).

In a more recent review Katzell (1974) concluded that job enrichment "works best for brighter, less alienated people."

Critical to most programs of job redesign is the alteration of one of Hackman and Lawler's (1971) four "core characteristics" of jobs--variety. The common assumptions are that increasing variety--through job enlargement, enrichment, or rotation--will reduce the monotony of jobs and that the ensuing reduction of boredom will lead to greater satisfaction and productivity. But this chain of causal assumptions has yet to be forged solidly, and the personal or environmental conditions that may limit the assumptions are scarcely understood. As Dickson has said, "job rotation and and job enlargement schemes enlarge the levels of variety . . . in individual jobs without any clear conception of the fundamental variables involved (1973, p. 732)."

The research reported here concerns three such variables:

1. Task characteristics, measured independently of the worker doing the task, with a particular emphasis on those task characteristics that provide variety in jobs.

2. Monotony, or the worker's perception of the amount of variety in his or her job.

3. Boredom, or the worker's affective reaction based on this perception. Excluded from the present research were considerations of the neurological, pharmocological, and behavioral concomitants of boredom.

The analyses were designed to answer three questions:

1. What is the contribution of task characteristics to the experience of occupational monotony and to feelings of boredom? On the basis of very early work by Florence (1924), Li (1928), Munsterberg (1913), and Wyatt, Langdon, and Stock (1937) and more recent laboratory work by Geiwitz (1966), it could be anticipated that the relationship would not be very great, since previous investigations had generally concluded that work lacking in variety was not always perceived as repetitive. It was also predicted that task characteristics would be better predictors of monotony than of boredom, the latter reaction being further removed from the environment in the assumed causal sequence.

2. What aspect or aspects of the temporal configuration of tasks were the most effective determinants of monotony and boredom? The three critical aspects investigated were the number of tasks done as part of the job, the lengths of these tasks, and how often each was repeated.

3. What individual differences moderate the relationship (a) between task characteristics and perceived monotony, and (b) between perceived monotony and feelings of boredom? Do individual differences affect workers' perceptions of the amount of variety provided by their work environments? Or do these differences influence their subsequent reactions once these perceptions have been established? Or both?

Studies of occupational boredom and tolerance of monotony have identified several personal characteristics that may ameliorate the effects of jobs that lack variety. These moderators are strikingly similar to those that circumscribe the success of job redesign programs. Geiwitz (1964) cites as relevant variables intelligence, creativity, desire for variety in social activity, and the tendency to daydream. Using previous research to draw a composite picture of "the personality of the person who is satisfied in doing repetitive work," Smith characterizes this person by

> . . . contentment with the existing state of affairs, placidity, and perhaps rigidity. His satisfaction would seem to be more a matter of close contact with and acceptance of reality than of stupidity or insensitivity (1955, p. 329).

On the basis of her own work, however, Smith was able to confirm only

> . . . that the susceptible worker is likely to be young, restless in his daily habits and leisure-time activities, and less satisfied with personal, home, and plant situations in aspects not

> directly concerned with uniformity or repetitiveness (1955, p. 329).

From this compendium of moderator variables two were selected for analysis: intelligence and off-the-job activity level.

Method

Subjects

The 370 workers studied were employed by two automotive suppliers, a printing company, and a research and development firm. Twenty-six percent were in managerial or professional occupations; 17 percent were clerical workers; and the remaining 57 percent were operatives, craftspeople, or service workers. The median age of the sample was 35 years, and their median years of schooling was 12.1. Fifty-four percent of the sample were women.

Procedures

Workers were first interviewed in their homes; a month later they were observed on their jobs on two separate occasions. The training of the observers, as well as the general observation methods used, have already been described in Chapter 4. Each observer devoted 20 minutes of each hour-long observational period to keeping a continuous record of the discrete tasks performed by the worker during that period. This task description information was subsequently coded in terms of total number of tasks performed and the frequency and duration of each. The analysis sample was confined to workers who were observed twice for exactly 20 minutes without being interrupted for reasons not related to their work.

Measures

Observed task characteristics. Five aspects of the observed tasks were used as predictors of monotony and boredom.

1. Number: the total number of different tasks performed by the worker.

2. Durations: the mean time of the tasks, computed over the total number of tasks that the worker performed.

3. Repetitions: the number of times that the most frequently repeated task was repeated.

4. Ratio of durations and repetitions. The duration of each task, in minutes, was first divided by the number of times the task was repeated. These ratios were then summed over all tasks performed.

5. Entropy. This measure, borrowed from information theory, had been used by Hill (1975) and Dickson (1973) to describe work-loads. The entropy of each task was $p \log_2 p$, where p was the proportion of the work period devoted to the task. The entropy of the total job was the sum of the entropies of its constituent tasks. This entropy measure tends to reduce the contributions of minor task variations to an observational measure of overall job variety. According to Dickson,

> . . . the work of one day could consist of packing one product for seven hours and six products for five minutes each. Alternatively, the work of one day could consist of packing seven different products for one hour each. In both of these days' work, the number of different tasks is seven and the number of changes of task is six, but the entropy measures of variety are higher for the second day's work (1973, p. 723).

Monotony. Two fixed-alternative questions were used to measure workers' perceptions of how monotonous their jobs were, one referring to how repetitive their tasks were and the other referring to how much variety there was. Observers used similar questions dealing with repetitiveness and variety to provide their subjective estimates of the monotony of the observed jobs.

Boredom. This measure used two fixed-alternative questions: "How often on your job does the time seem to drag?" and "How interesting is your work?"

Intelligence was estimated from interviewers' ratings of workers. In a related study (Quinn & Shepard, 1974) the same interviewers' ratings correlated .72 with tested intelligence as measured by the Amons Quick Test.

Off-the-job activity level. This measure was the amount of the worker's participation in 20 recreational, social, and familial activities not related to work.

Results

The product-moment correlations between task characteristics and measures of monotony and boredom appear in Table 1. Since differences between these Pearson r's and comparable eta coefficients of association were small, it could safely be concluded that the observed relationships were generally linear.

Task characteristics expressed in various ways were significantly related both to the perception of monotony and to feelings of boredom.

Table 1

Correlations (Pearson r's and Etas)[a] between Task Characteristics, Perceived Monotony, and Feelings of Boredom (N ~370)

Task characteristics[b]	Perceived Monotony		Worker's feelings of being bored
	By worker	By observer	
Number	.07 (.06)	.36** (.35)	.13* (.11)
Durations	.22** (.27)	.30** (.43)	.11* (.17)
Repetitions	.26** (.27)	.58** (.58)	.19** (.21)
Ratio of durations and repetitions	.20** (.28)	.36** (.47)	.10 (.19)
Entropy	.20** (.24)	.24** (.33)	.08 (.10)

*p < .05.

**p < .01.

[a] Eta coefficients are shown parenthetically. In computation of etas, each task characteristic was trichotomized, and the monotony and boredom measures were treated as continuous variables.

[b] For the computation of Pearson r's the directions of some measures of task characteristics were reversed so that a higher score on a measure would always represent greater observed monotony. All Pearson r's should therefore be expected to be positive.

They were, predictably, most strongly related to the observers' perceptions of monotony in the jobs they observed, somewhat less strongly related to workers' own perceptions of monotony in their jobs, and least strongly related to workers' feelings of boredom.

The poorest observational predictor of monotony and boredom was the number of tasks performed during the observational periods. The best and most parsimonious predictor was the number of times the most frequently-repeated task was repeated, with the duration of the average task a close second. Neither of the conceptually and mathematically more ornate estimators of task variety--that is, entropy and the ratio of durations and repetitions--was superior to the simple measure of repetitions as a predictor of perceived monotony. Nor was either significantly related to workers' feelings of boredom.

The moderating effects of intelligence and off-the-job activity level are indicated by the differences between the pairs of correlations in Table 2. Since repetitions were the best and simplest predictors of monotony and boredom, the task characteristics referred to in Table 2 are based only on the repetitions measure.

Intelligence had no significant effect upon the association between job characteristics and either monotony or boredom. The moderating effects of intelligence were confined to the relationship between perceived monotony and feelings of boredom. That is, monotony was more closely associated with boredom among more intelligent workers. The moderating effects of off-the-job activity were likewise strongest on the relationship between perceptions of monotony and feelings of boredom,

Table 2

Correlations (Pearson r's) between Monotony and Boredom for Workers Differentiated by Intelligence and Off-the-Job Activity Level

Indicator of monotony and boredom	Intelligence			Off-the-job activity level		
	High (N=168)	Low (N=202)	Difference	High (N=174)	Low (N=196)	Difference
Observed variety (repetitions), by . .						
..Observers' perceptions of monotony	.56	.57	-.01	.61	.54	.07
..Workers' perceptions of monotony	.34	.16	.18	.25	.23	.02
..Workers' feelings of being bored	.27	.12	.15	.30	.09	.21*
Workers' perceptions of monotony, by workers' feelings of boredom	.52	.31	.21*	.59	.17	.42**

*$p < .05$.

**$p < .01$.

[a] Some measures were reversed so that a higher score would always reflect greater monotony or boredom. All correlations should therefore be expected to be positive.

with the latter two variables being more strongly related among workers who were more active off their jobs.

Discussion

That the associations between task repetitiveness and both monotony and boredom were palpable but small suggests that the major sources of occupational boredom and monotony may possibly be sought in working conditions other than task variety.

Some leads in this search are provided in Fenichel's (1951) psychoanalytic approach to boredom and the neurophysiological ones of Berlyne (1960) and Hebb (1958). Geiwitz's experimental tests and synthesis of these approaches identify cognitive arousal and environmental constraint as the two major determinants of boredom, relegating repetitiveness to an equivocal position. Job enlargement programs often involve many simultaneous organizational and task changes. Geiwitz's theory would suggest that what should be altered in such programs should, oddly enough, _not_ be variety, either exclusively or even principally. Instead, it should be yet another of Hackman and Lawler's core dimensions--autonomy. Even matters commonly dismissed as palliatives in job redesign--such as music at the workplace--have a firm psychological base as possible reducers of boredom in Geiwitz's theory.

Among the task characteristics capable of predicting monotony and boredom, the best was how often the most-frequently-repeated task was repeated. This would suggest that task redesign efforts will be unsuccessful if they are confined simply to grafting on a few peripheral activities "for variety's sake" to a basically highly repetitive job. The unaltered repetitions would defeat such a cosmetic effort.

References

Berlyne, D. Conflict, arousal, and curiosity. New York: McGraw-Hill, 1960.

Dickson, J. The physical correlates of variety in work. Human Relations, 1973, 26, 715-734.

Fenichel, O. On the psychology of boredom. In D. Rapaport (Ed.), Organization and pathology of thought. New York: Columbia University Press, 1951.

Florence, P. Economics of fatigue and unrest and the efficiency of labor in England and American industry. New York: Holt, 1924.

Geiwitz, P. The structure of boredom. Doctoral dissertation, The University of Michigan. University Microfilm No. 65-5305, 1964.

Geiwitz, P. Structure of boredom. Journal of Personality and Social Psychology, 1966, 3, 592-600.

Hackman, J., & Lawler, E. E., III. Employee reactions to job characteristics. Journal of Applied Psychology, 1971, 259-286.

Hebb, D. A textbook of psychology. Philadelphia: Saunders, 1958.

Hill, A. B. Work variety and individual differences in occupational boredom. Journal of Applied Psychology, 1975, 60(1), 128-131.

Hulin, C., & Blood, M. Job enlargement, individual differences, and worker responses. In L. Cummings and W. Scott, Jr. (Eds.), Reading in organizational behavior and human performance. Homewood, Ill.: Irwin Dorsey, 1969.

Katzell, R. Structural approaches to increasing both job satisfaction and productivity. Paper presented at the 82nd annual meeting of the American Psychological Association, New Orleans, 1974.

Lawler, E. E., III. Job design and employee motivation. Personnel Psychology, 1969, 22, 426-435.

Li, C. A summer in the Ford works. Personnel Journal, 1928, 7, 18-32.

Munsterberg, H. Psychology and industrial efficiency. Cambridge: Houghton-Mifflin, 1913.

Quinn, R., & Shepard, L. The 1972-73 quality of employment survey. Ann Arbor, Michgian: Survey Research Center, 1974.

References (continued)

Quinn, R., Staines, G., & McCullough, M. Job satisfaction: Is there a trend? Washington, D. C.: U. S. Department of Labor, 1974.

Smith, P. The prediction of individual differences in susceptibility to industrial monotony. Journal of Applied Psychology, 1955, 39, 322-329.

Wyatt, S., Langdon, J., & Stock, F. Fatigue and boredom in repetitive work. Industrial Health Research Board, London, Report No. 77, 1937.

Chapter 15

PREDICTING WITHDRAWAL FROM WORK: ABSENTEEISM AND TURNOVER

by

Nina Gupta and Terry A. Beehr

ABSTRACT

This inquiry is concerned with two common and costly forms of withdrawal from work: absenteeism and voluntary resignation. A primary aim was to extend the search for motivational causes and situational conditions giving rise to withdrawal. A secondary aim was to see whether absence and quitting are additive forms of a general withdrawal syndrome. The main focus was on work role stress as a source of motivations to withdraw. Individual differences in stress tolerance were examined. Differences in demography and in life context were examined to see how they might operate separately, or in conjunction with motivational factors, in withdrawal behavior.

Four role stresses were treated: role ambiguity, role overload, underutilization of skills, and inadequacy of resources for work performance. Two forms of withdrawal were measured, each with more than one method of measurement. The demographic and background factors treated were: age, sex, marital status, family composition as to children, job tenure, and the presence of problems with transportation and work hours.

The results show that all four role stresses are associated with absenteeism and turnover. Elevated absenteeism tends to precede quitting. There is some support for the idea that individual tolerance for stresses may moderate their effects. Role stresses appear to be somewhat more strongly associated with later absenteeism than with prior absenteeism, suggesting that the stresses are causal. The demographic and life context variables were as effective in explaining withdrawal behavior as were the role stresses, confirming that absence and quitting are in part determined by risks, opportunities and constraints that are unrelated to the characteristics of the job itself. However, work role stresses have their own causal effects even when the effects of demography and background are controlled. Overall, the measured effects are rather weak, and one must conclude that simple and general models of causation are not adequate. It is likely that a fuller explanation will require treatment of interaction patterns among the personal, job, demographic and contextual factors here shown to be relevant.

Chapter 15

PREDICTING WITHDRAWAL FROM WORK: ABSENTEEISM AND TURNOVER

Absenteeism and turnover have been the subject of much inquiry in search for understanding of the causes and assessment of the costs and consequences. Interest continues because they are still not well understood, and because they are known to be forms of member behavior that are troublesome and costly for employing organizations and often associated with personal disadvantage or stress for the individuals involved. Another reason for continuing interest is the realization that absences and job turnover are related in some way to the individual's need for accommodation to life stresses and opportunities, and society's need for some adaptive fluidity in the labor market.

The present study extends and replicates the prior research in two directions: (1) The prediction of these behaviors from attributes of the person, the work situation and the individual's life situation, and (2) The determination of the relationship between the two forms of withdrawal--i.e., whether they are sequentially related, or independent alternative forms of withdrawal, or causally unrelated.

Prior Research: Prediction of Withdrawal

A number of plausible factors have been employed in the prediction of absenteeism and turnover. These include demographic and life-background characteristics of individuals (e.g., Baumgartel & Sobol, 1959; Boyd, 1961; Saleh et al., 1965; Farris, 1971); individual personality and attitude characteristics (e.g., Lyons, 1971); job characteristics (e.g., Guest, 1955; Turner & Lawrence, 1965;, Katzell, 1968; Hackman & Lawler, 1971); the work environment (e.g., Hewitt & Parfitt, 1953; Evans 1963); organizational characteristics (e.g., Lawler & Hackman, 1969; Ingham, 1970; Schefler et al., 1971); and "external" conditions such as the state of the labor market (Behrend, 1953). A review of these and other prior inquiries suggests three methodological and conceptual problems of particular interest.

Unclear causality. Nearly all of the studies mentioned were concurrent rather than longitudinal in design intention, thus raising the issue of direction of causality when, for example, an attitude such as dissatisfaction is found to be related to absenteeism or decision to quit. To compound the difficulty, many of the studies were not truly concurrent, but employed measures of absenteeism for a prior period of time, or measured individual attributes, perceptions and attitudes during exit interviews after the decision to leave the job. While some predictors such as age, educational attainment, job tenure, are surely impervious to reverse causation, one can not be confident that the working conditions experienced by the individual remained constant during a prior period encompassing the withdrawal behavior, nor can one be confident that the attitudes upon departure are those that prevailed prior to the decision to quit. In short, many of the studies were not even truly concurrent, but actually reversed the logically required conditions for inferring causality from an order of occurrence in time. Since absences and quits are episodic behaviors rather than stable, continuous behaviors, it is particularly important that the investigation of precursors should rest upon their measurement prior to the observation of the behaviors. Porter and Steers (1973) note that the more recent studies of withdrawal behavior are more often free of this methodological weakness.

Joint use of predictors. Investigators of withdrawal behavior have tended to confine their approach to one or another class of predictors and to neglect to treat diverse predictors jointly. A common type of psychologically oriented study, for example, may successfully predict withdrawal behavior from job satisfaction, but without reference to the demographic characteristics that may be the source of both dissatisfaction and withdrawal, a point noted in several reviews of the field (Brayfield & Crockett, 1955; Vroom, 1964; Porter & Steers, 1973). On the other hand, many studies attempt to predict withdrawal exclusively from demographic data (Parnes et al., 1971; Kim, Rodrick, & Shea, 1973; Rodrick & Davis, 1973). Some studies employ a diverse array of predictors -- demographic, environmental, and psychological -- but neglect to treat their interactions or combined effects (Quinn, Fine, & Levitin, 1970; Farris, 1971). It seems likely that withdrawal behaviors are complexly determined, and not to be understood except through complex sets of predictor variables.

Global predictors and proxy predictors. An understanding of withdrawal behaviors is more likely to be achieved if the predictors examined are specific rather than composite in nature, and if generic variables describing persons or their environments are used instead of convenient substitutes. The latter point is best illustrated by the common use of demographic predictors, such as age or sex, which commonly "predict" withdrawal but leave open for speculation the interpretation of that result. Is it age or sex, as such, that is the operative cause, or perhaps some age-related psychological change, or some feature of contemporary sex role prescriptions? A commonly used global predictor is an index of overall job satisfaction (e.g., Kerr, Koppelmeier & Sullivan, 1951). While it is helpful to establish that job satisfaction is a significant correlate of withdrawal, it would be more helpful to know what more specific domains or facets of satisfaction are at work. Studies by Hulin (1968) and Mangione (1972) show that the sub-components of satisfaction and of working conditions operate differently in relation to withdrawal.

Overview and the Present Study

The study reported here was designed to minimize the methodological and conceptual issues mentioned above. The data were longitudinal, with measurement of the selected predictors prior to the occurrence of the withdrawal behaviors of absence and turnover. Job related stress was advanced as a potential cause of withdrawal, and stress was measured not globally, but with respect to four differentiated types of stress. Background and demographic variables were employed jointly with other predictors. Also, invoking the idea of person-job fit as a relevant consideration, use is made of measured individual differences in tolerance for the presumed stressful job attributes.

Role Stress

While evidence bearing directly on the relationship between role stress and withdrawal is relatively sparse, it suggests a positive association between the two. Weitz (1956), for example, found that role ambiguity (a role stress) was related to continued participation in an organization. Similarily, Lyons (1971) found perceived role clarity to be negatively related to both intention to leave, and actual voluntary turnover. Taylor & Weiss (1969, cited in Porter & Steers, 1973) reported that perceived opportunities to use one's skills and abilities were inversely related to turnover. Skill underutilization can be considered a role stress.

The studies mentioned above used turnover, rather than absenteeism, as the criterion variable. The arguments that can be made, however, with respect to the relationship of stress to either absenteeism or turnover, are similar if both absenteeism and turnover represent alternative degrees or forms of withdrawal from an employing establishment. One may reason that withdrawal is induced because the stressful work situation is unpleasant. Effects of stress would, therefore, result in chronic absenteeism and, if the situation were aversive enough, termination.

In addition to this direct effect of stress on withdrawal, stress may have indirect effects on absenteeism and turnover because of the relationship of both of these variables with job satisfaction. For example, it has been demonstrated that role clarity is related to high satisfaction (Lyons, 1971; Beehr, 1974). Studies cited earlier showed that job dissatisfaction results in high absenteeism and turnover. Stress may, therefore, have an effect on absenteeism through the mediation of job satisfaction.

Background Characteristics

The effects of background characteristics on withdrawal behaviors have been investigated at length in past research (Schuh, 1967). The impact of these background characteristics on absenteeism and turnover are, however, not always the same in all situations. Consistent relationships have been found, for example, between age and turnover (e.g., Ley, 1966; Robinson, 1972), but the relationship between age and absenteeism is not equally clear cut. In addition, while age is inversely related to turnover, the relationship between age and absenteeism is generally positive or zero, rarely negative (e.g., Cooper & Payne, 1965).

A similar conclusion may be drawn with respect to the relationship of organizational tenure to withdrawal. While employees with high tenure appear less likely to terminate, the data are not consistent with respect to absenteeism. While some studies have found no relationship between tenure and absenteeism (e.g., Hill & Trist, 1955), others have found different patterns of relationship among different demographically defined subgroups (e.g., Baumgartel & Sobol, 1959).

Family size is another variable that has been found to be related to withdrawal behaviors. The sex of the employee has similarly been found to affect withdrawal tendencies.

The above variables were incorporated in the present inquiry along with two others that appeared, *prima facie*, to be potentially influential: (1) Problems of work schedule/hours, and (2) Problems of travel to and from work.

Tolerance for Role Stress

There have been few studies that explicitly include measures of individual differences in stress tolerance. An exception is that of Lyons (1971) who found that low role clarity was associated with withdrawal, but only for employees high in need for clarity. It appears likely that the effect of role stress on withdrawal will be minimal for those having high tolerance for stress. The present study examines this possibility with reference to four potentially stressful facets of work roles.

Predicting Withdrawal

The central thrust of this inquiry is defined by the foregoing discussion. It was proposed that work role stress (four distinguished forms of stress) can be shown to have a causal relationship to subsequent withdrawal behavior, and that these relationships will be: (1) Stronger among those individuals having relatively low tolerance for the respective stresses; (2) Not an artifact arising from demographic and life-situation variables.

It is also of interest to consider the span of time over which these proposed relationships may become stronger or weaker, but no expectations could be formulated on this matter. It is plausible that the associations between stresses and absenteeism might be relatively strong concurrently but diminish over time if the levels of stress remain constant. It is equally plausible, however, that the impact of stresses upon absenteeism and turnover might take some time to develop, thus showing a lag effect. Similarly, it is of interest that the four work role stresses may be differential in their potency to induce withdrawal, but in the absence of theory and prior evidence, no predictions were made.

Relationship between Absenteeism and Turnover

To say that absenteeism and turnover are both withdrawal behaviors leaves ill-defined the nature of the relationship between them. Some controversy continues around the issue of the existence and the type of relationship between absenteeism and turnover. Three alternative forms of relationship have been proposed (Beehr & Gupta, 1978).

The first position argues that absenteeism is a lesser and earlier form in a withdrawal behavior cycle that is completed when an employee quits the job (e.g., Herzberg et al., 1957). There is an implication that the same causes and predictors apply to both. Melbin (1961) concluded that "High absenteeism . . . appears to be an earlier sign, and turnover . . . the dying stage, of a long and lively process of leaving" (p. 15).

The second position argues that absenteeism is an alternative to quitting (e.g., Hill & Trist, 1955). In other words, some employees may respond to particular psychological and environmental factors by staying in the organization but exhibiting chronic absenteeism, while others respond to the same conditions by quitting the organization. In this view, the psychological dynamics are the same, but absenteeism need not necessarily be a lesser response or precursor to ultimate turnover.

The third and similar position, advocated by March and Simon (1958), argues that there is no consistent relationship between absenteeism and turnover, as "differences between absences and turnover stem not from differences in the factors inducing the initial impulse but primarily from differences in the consequences of the alternative forms of withdrawal." That is, it is opportunity, cost, or convenience, not individualistic psychodynamics, that determines the form of withdrawal.

From a review of the literature in this area, Lyons (1972) concluded that the evidence with respect to the relationship between absenteeism and turnover measured at the individual level is consistent: there is a significant and positive relationship between the two forms of withdrawal. He concluded further that there is tentative support for the notion that there is a progression of withdrawal, absenteeism being the earlier and lesser form. He concluded finally that there was little evidence to support or deny March and Simon's (1958) assertion of common correlates and opportunistic forms of withdrawal.

The study reported here allowed some clarification of these alternative and complementary views, in three ways: (1) Replication and confirmation of the positive relationship between absenteeism and turnover at the individual level; (2) Comparison of absenteeism and turnover as to their having the same, or different, predictors and imputed causes; (3) Testing of the sequential hypothesis that absenteeism is an intervening response to job stress, leading to later turnover rather than substituting for turnover.

Definition and Measurement

With respect to absenteeism, a choice is necessary between a definition that emphasizes the amount (or proportion) of scheduled time missed from work or instead emphasizes the rate or frequency of occurrence of absences. These alternative approaches to definition and measurement, of course, produce measures that are strongly correlated, but there is evidence that they have significantly different properties (Metzner & Mann, 1953; White, 1960; Huse & Taylor, 1962; Smith, 1976). The former (amount of time missed) captures absenteeism more usefully when the interest is specifically in the management policy and cost implications; the latter appears to measure absenteeism more usefully when the interest is in the factors associated with the individual's decision to be absent. The present study is of the latter kind, so the definition adopted for measurement purposes refers to the initiation of an episode of absence without regard to its duration.

A similar rationale was applied to the distinction between voluntary and involuntary absences. Since the focus of the study was upon motivations to withdraw, it seemed best to omit involuntary absences to the extent possible. Lyons' (1972) review of the research literature confirms that excused and unexcused absences correlate differently with characteristics of individuals and of organizations. Macy and Mirvis (1976) note that involuntary absences, as in the case of a long-term illness, are less likely than voluntary absences to reflect intentional or unconscious withdrawal processes. While the distinction between voluntary and involuntary absences presented some difficulties in actual measurement, an effort was made to include only absences within the control of the individual and not scheduled in advance with the employer.

Similar issues of definition arise with respect to turnover. From the employer's point of view an event of departure of an employee from his job is viewed differently depending upon whether the departure appears to imply a loss or a gain for the employer. A change to another job within the firm as in the case of a promotion or requested transfer may have accommodative value, as may a discharge initiated by the employer, while a quit initiated by the employee is likely to imply a loss. The present study aimed to treat the motivational and situational forces acting upon the individual to cause withdrawal behavior. The definition employed, following the rationale offered by Price (1972) and Macy and Mirvis (1976), was to include only instances of turnover that were voluntary and involved leaving the employing organization. Such a neat definition, however, is not easily applied in practice. While departure for reasons of permanent layoff, death, or retirement seem to be clearly involuntary, a record of discharge at the initiative of the employer may often conceal a process of separation in which both parties are involved and in which the employee may prove the true initiator. Quinn, Fine, and Levitin (1970) noted some "quits" that in fact were stimulated by the employer, and a high proportion of "discharges" that were merely the delayed official recognition that the employee voluntarily was no longer coming to work or not reliable in attendance. Ideally, records of different employers would be uniformly kept, and the views of both parties could be employed in assessing the instance of turnover. In this study, a departure was assumed to have a voluntary component unless the employer's record keeping practices provided a fairly strong presumption to the contrary.

Populations and Measures

Sample

Six hundred fifty one employees from five organizations participated in the first phase of the study. These organizations were a printing company (N=173), a small research and development company (N = 21), two automotive supply companies (N's = 120 and 124), and four service departments of a hospital (N=213). The response rate was 72.9 percent. Employees from three of the five organizations were recontacted for Phase II of the study. Two hundred seventy two agreed to participate. Interview data were collected on two occasions, approximately 20 months apart. Further details on the populations studied are provided in Chapters 1 and 2 of this volume.

Job stress. Interview measures on four stresses were obtained in Phase I of the study. Role ambiguity was measured using a four-item index described in Beehr (1976); role overload (two items) and under-utilization of skills (three items) were measured through indices described in Beehr (1974). Resource inadequacy stress was measured through the resource adequacy index (reversed) of the Quality of Employment Survey (Quinn & Shepard, 1974). Table 1 shows the stress measure intercorrelations.

Turnover. Four measures of turnover were used. Two of these constitute what better might be termed "turnover intent." In both phases of the study respondents answered the question:

> "Taking everything into consideration, how likely is it that you will make a genuine effort to find a new job with another employer within the next year -- very likely, somewhat likely, or not at all likely?"

Two additional measures of turnover were based upon search of the employers' records. One represented the elapsed time in months between the Phase I interview and the date of turnover, for those who had in fact left their original, Phase I, job. The second classified all respondents into four categories: (1) Remained in the same job; (2) Voluntarily left job and employer; (3) Left job but not employer -- e.g., transfers and promotions; and (4) Involuntarily left job and employer -- e.g., retirement, discharge, permanent layoff, etc. The present analysis refers only to categories 1 and 2. A departure was presumed to be voluntary unless there was evidence to the contrary. Table 2 shows selected intercorrelations.

Absence. Information on employee absences was obtained using both self-report in interview and records data. The interview measures of absence were obtained in both phases of the study, and consisted of three items. These three items were combined to form a single index here called

Table 1

Correlations among Role Stress Measures
(N ranges from 629 to 640)

	1	2	3	4
1. Role ambiguity	(.65)[a]			
2. Role overload	.47**	(.44)		
3. Skill underutilization	.07	.05		
4. Resource inadequacy	.42**	.33**	.10*	(.55)

* P < .05

** P < .01

[a] Diagonal shows internal consistency estimates (Alpha) where available.

Table 2

Correlations among Selected Indicators of Absenteeism and Turnover

(N's range from 217 to 614)

	1	2	3	4
1. Turnover intent Time 1	--			
2. Actual voluntary turnover, $T_1 - T_2$	.39**	--		
3. Absenteeism - Self-Report, Time 1	.09*	.20**	--	
4. Absenteeism - Recorded month prior to Time 1	.17**	.16*	.38**	--
5. Absenteeism - Recorded six months following Time 1	.23**	.30**	.62**	.46**

*p < .05

**p < .01

"reported absenteeism".

"Aside from vacation and holidays, how many days of scheduled work have you missed in the past month?"

"How many of these days did you miss just because you didn't feel like going to work that day?"

"Would you say you are absent more often than other people you work with, less often, or about the same?"

Data on the frequency of each employee's absences (excluding holidays) during the month prior to the Phase I interview were obtained from employers' records for all respondents interviewed in Phase I. Data on the frequency of absence for 23 months subsequent to the Phase I interview were obtained from employers' records for employees resurveyed in Phase II. One continuous period of absence regardless of duration was counted as one episode. Because the record-keeping, definition, and classification systems varied among the organizations, some uncontrolled differences exist on the recorded absenteeism variables. These differences, however, did not appear to be associated with distinctive patterns of correlational statistics, and it was judged appropriate to treat them as random measurement error. The following categories were counted in the recorded absenteeism variables for each of the organizations:

Printer:	Unexcused absences; excused absences attributed to illness.
R&D Firm:	All absences.
Auto Supplier I:	Unexcused absences; excused absences attributed to illness, accident, injury, and funerals.
Auto Supplier II:	Unexcused absences; excused absences attributed to illness or transportation failure.
Hospital:	Unexcused absences; excused absences attributed to illness.

For the population as a whole, absence behavior proved to be highly reliable in the sense that an individual is likely to display consistent frequencies across a series of sequential time periods, (Alpha coefficient .87) for one-month periods. Table 2 has further information on absence measures.

Demographic variables. During the Phase I interviews there were obtained several demographic descriptors employed in the present analysis, namely, sex, age at last birthday, marital status, and number of dependent pre-school and school age children in the respondent's household.

Problems with working hours. Respondents were asked, "Could you tell us what problems or difficulties you run into concerning the hours you work, your work schedule, or overtime?" If problems of any kind were mentioned, there followed: "How much of a problem for you is this (are these things)?"

Transportation problems. Respondents were asked, "What things concerning travel to and from work do you consider problems and would like to see changed if possible?". The follow-up question and coding procedures were the same as for the "hours" measure, above.

Tenure. Information on organizational tenure was obtained from personnel records. The month and year the respondent first started working for the organization were recorded. These data were transformed to a range from 1 to 41, representing the number of years the respondent had been with the organization.

Stress tolerance. Three areas of employee preference were measured in the interview. The measure for tolerance of ambiguity was the mean of three true/false items (Alpha Coefficient .42):

> A good job is one where what is to be done and how it is to be done are always clear;

An expert who doesn't come up with a definite answer probably doesn't know too much;

and,

What we are used to is always preferable to what is unfamiliar.

The measure of tolerance for underutilization of skills was the mean of three four-point scales, with responses ranging from "not at all desirable" (1) to "very desirable" (4). The items were (Alpha Coefficient .71):

How desirable to you is it that . . .

. . . you have a chance to develop your own special abilities at work?

. . . you get a sense of accomplishment from your work?

. . . your work is challenging to you?

The measure of tolerance for resource inadequacy was also the mean of three four-point scales with the same response format as the previous set. The items were (Alpha Coefficient .61):

How desirable to you is it that . . .

. . . the people you work with are helpful?

. . . your supervisor is helpful?

. . . you receive enough help and equipment to get your job done?

The interview did not contain any measure relevant to tolerance of role overload.

Results and Discussion: Prediction of Withdrawal

The analytical results will be displayed in five segments treating, respectively: Role stress as a cause of absenteeism; Role stress as a cause of turnover; The moderating effects of stress tolerance; The implications of certain demographic and background variables; and, The nature of the relationship between absenteeism and turnover.

Stress and Absenteeism

The exploration of the causal role of job stress in relation to absenteeism rested upon measures of four classes of work role stress at time 1 and actual recorded voluntary absence rates for various periods subsequent to time 1. For convenience and economy of display, Table 3 shows the results by month for the first six months and for two multi-month summary periods thereafter, these summary periods being chosen somewhat arbitrarily to match peak times of population loss due to turnover. The four work role stresses refer to ambiguity, overload, underutilization of skills, and resource inadequacy.

The overall results show relatively weak correlations between the respective stresses and absenteeism, with none greater than r=.21. However, a substantial proportion of these correlations are statistically significant. All of the significant correlations but one are positive, as predicted, and there appears to be some patterning of these results as to type of stress and period of time. Each of the four stresses shows some significant correlations with absenteeism, and in the case of one stress, resource adequacy, the relationship persists beyond the period of six months. The strength of the relationships appears to diminish with time, and in one case role ambiguity unaccountably becomes negative. Underutilization of skills appears to lose its association with absenteeism earlier than the other stresses. It can be speculated that these anomalous results might be a consequence of selective population loss over the period, the turnover rate being particularly high, for example, among those initially reporting

Table 3

Correlations of Stress Measured at Time I with Recorded Absenteeism Subsequent to Time I (Ns in parentheses)

	Reported Stress at Time I			
Absenteeism	Role Ambiguity	Role Overload	Under-Utilization of Skills	Resource Inadequacy
Month after Time I:				
First	-.04 (346)	.06 (341)	.02 (346)	.01 (340)
Second	.15** (341)	.15** (336)	.15** (341)	.14* (325)
Third	.08 (330)	.11* (325)	.16** (331)	.13* (324)
Fourth	.15** (329)	.12* (324)	.06 (330)	.16** (323)
Fifth	.16** (324)	.16** (319)	.00 (325)	.17** (318)
Sixth	.04 (304)	.08 (300)	.01 (305)	.09 (298)
Summary Indices (Months after Time I):				
First Six	.13* (329)	.16** (324)	.10 (330)	.16** (323)
Seventh - Nineteenth	-.01 (242)	.09 (240)	.03 (243)	.21** (237)
Twentieth - Twenty third	-.17* (136)	.02 (136)	-.06 (137)	.07 (136)

*$p < .05$
**$p < .01$

underutilization of skills. In summary, the data provide evidence confirming the hypothesized association between work role stresses and absenteeism.

The role of stress as a cause of absenteeism, rather than a consequence or simple correlate, would be supported if the concurrent (same month) relationships between stress and absenteeism proved to be weaker than the relationships between stress at Time 1 and absenteeism during following months. Table 4 shows the outcome of such an analysis, using only those individuals for whom complete data sets were available for the period. For all four stresses, the correlations are indeed greater with reference to the following months than for the month concurrent with the measurement of stress. In three of the four tests, the difference between correlations is sufficient to reach statistical significance. The differences are not an artifact arising from comparing a one-month absenteeism rate with a six-month rate. While the analysis is made ambiguous as to causation because of lack of control over possible changes in stress levels over the period, the evidence at hand appears to support the conclusion that work role stresses play a causal or at least antecedent role in inducing absenteeism.

Stress and Turnover

Table 5 displays the relationships between the four work role stresses as measured at Time 1 and four indicators of turnover or intent to leave. The concurrent measures of stress and of intent to leave (Phase I measurements) show a consistent result: all four stresses are positively and significantly related to intent to leave. This result might have been replicated with respect to later intent to leave (Time 2) except for the effects of the passage of time and the selective loss of individuals from the original population. In the intervening months the available population

Table 4

Comparison of Correlations between Stress and Subsequent Absenteeism with Correlations between Stress and Concurrent Absenteeism
(N=299)

	Reported Stress at Time 1			
Recorded Absenteeism	Role Ambiguity	Role Overload	Under-Utilization of Skills	Resource Inadequacy
	**[a]	*		**
Month Prior to Time 1 Interview	-.08	.03	.05	-.03
First Six Months after Time 1 Interview	.14*	.14*	.11*	.14*

*p < .05
**p < .01
[a] The asterisks next to the boxes indicate the significance of the differences between correlations in the box, using one-tailed tests.

Table 5

Correlations of Stress Measured at Time 1 with Reported Turnover Intent and Actual Turnover

	Reported Stress at Time 1			
Turnover	Role Ambiguity	Role Overload	Under-Utilization of Skills	Resource Inadequacy
Turnover Intent at Time 1 (N = 620)	.13**	.12**	.29**	.15**
Turnover Intent at Time 2 (N = 154)	.10	.03	.04	.18*
Voluntary Turnover (N = 276)	-.04	.07	.23*	-.04
Elapsed Time to Voluntary Turnover (N = 47)	.05	-.02	.17	.11

*p < .05
**p < .01

dropped from 620 to 154, in part for unbiased reasons (e.g., not asked to participate in the second phase of measurement) but also because of the departure of 79 individuals from their original jobs.[1] In any case, the relationship between stresses at Time 1 and intent to leave diminished in strength, with only one stress, resource inadequacy, retaining a significant relationship. The correlations between stress at Time 1 and actual voluntary turnover were insignificant, except for one stress, underutilization of skill. Stress at Time 1 appears to be unrelated to the duration of elapsed time before departure. In summary, work role stress is clearly associated with concurrent intent to leave, but only weakly if at all to actual subsequent voluntary departure. Underutilization of skill appears to be the only work role stress associated with actual decision to leave.

Moderating Effects of Stress Tolerance

If work role stress is considered to be an antecedent, and possibly a prime cause, of withdrawal behavior, then one should expect that these relationships would be diminished for individuals having a high tolerance for the stresses and exaggerated for those low in stress-specific tolerance. This hypothesis could be tested with reference to only three of the available work role stress measures, omitting role overload. The results are displayed in Table 6 with reference to absenteeism, and in Table 7 with reference to turnover. In each case, the population was dichotomized at the

---

[1] Of the 79 job leavers, 50 left voluntarily, 5 were recorded as dismissed for cause, and the remaining cases were attributed to death, disability, permanent layoff, retirement. An additional number of the original cases were "lost" to this analysis by promotion or transfer within the same employing organization.

Table 6

Correlations of Role Stress and Absenteeism for Employees with Low versus High Tolerance of Stress (N's in parentheses)

Reported Stress at Time 1	Degree of Tolerance for the Stress	Time 1 Reported Absenteeism	Time 1 Recorded Absenteeism	Time 2 Reported Absenteeism	Time 2 Recorded Absenteeism 2nd Month	Time 2 Recorded Absenteeism 1st Six Months
Role Ambiguity	High	-.06 (139)	.03 (89)	-.09 (39)	.06 (53)	.12 (51)
	Low	.09 (492)	.04 (433)	-.20* (115)	.16** (282)	.12 (272)
Under-utilization of Skills	High	.11 (290)	.03 (262)	-.14 *[a] (71)	.19* (176)	.13 (174)
	Low	.15** (333)	.06 (257)	.19 (83)	.14 (157)	.12 (148)
Resource Adequacy	High	.09 (298)	.08 (239)	.16 (75)	.09 (164)	.10 (159)
	Low	.10 (325)	.01 (278)	-.11 (77)	.19* (168)	.22** (161)

*p < .05

**p < .01

[a] The asterisk next to the box indicates the significance of the difference between the correlations in the box, using a one-tailed test.

Table 7

Correlations of Role Stress and Turnover for Employees with Low versus High Tolerance of Stress
(N's in parentheses)

Reported Stress in Time 1	Degree of Tolerance for the Stress	Correlations with Turnover: Time 1 Turnover Intent	Time 2 Turnover Intent	Voluntary Turnover	Time to Voluntary Turnover
Role Ambiguity	High	.12 (139)	.19 (39)	.04 (73)	-.09 (13)
	Low	.14** (520)	.03 (114)	-.06 (206)	.04 (34)
Under-utilization of Skills	High	.26** (294)	.03 (71)	.21* (125)	.33 (20)
	Low	.30** (337)	.02 (82)	.23** (152)	.03 (27)
Resource Inadequacy	High	.11* (304)	-.06 **[a] (75)	-.09 (144)	-.01 (26)
	Low	.20** (328)	.41** (76)	.01 (131)	.25 (21)

*$p < .05$
**$p < .01$
[a] The asterisk next to the box indicates the significance of the difference between the correlations in the box, using a one-tailed test.

population median score on each of the stress-tolerance measures. Correlations between stress exposure and withdrawal were then calculated separately for the high-tolerance and low-tolerance segments of the populations.

Table 6, absenteeism, shows that in ten out of the fifteen paired correlations, the correlation for the low-tolerance individuals is the greater, but in only one instance does the difference reach significance. Similarly, in Table 7, turnover, eight of the twelve correlation pairs show the one for low-tolerance individuals to be the greater, but in only one instance does the difference reach significance. In both tables there is a slight tendency for significant correlations to occur more frequently in the low-tolerance groups.

It must be concluded that the effects of work role stress upon absenteeism and turnover appear with some consistency to be greater among individuals with low specific stress tolerance, but that a moderating role for stress tolerance is very slight, at best, and not proven by these results.

Demographic and Life Context Factors

It was argued earlier that attempts to assess the generic causes of withdrawal behavior, such as work role stress, are often clouded by the presence of unknown factors of population composition that may work either to exaggerate or to conceal the true causal relationships. The analyses to be summarized here were designed to test the hypothesis that variations in withdrawal behavior (absenteeism and turnover) within non-homogeneous populations are significantly associated with work role stresses after the removal of possibly confounding effects arising from demographic and life context variables. The analytic procedures for testing the proposition involve, first, determining the combined effects of significant background factors upon the withdrawal behavior, and second, determining the effects of work role stresses upon the withdrawal behavior after the withdrawal measures have been adjusted to remove the effects of background factors.

The data with respect to absenteeism are shown in Tables 8 and 9. Table 8 reports the beta-coefficients for each of three background variables in relation to each of the four absenteeism measures used in prior analyses. Background variables having no significant implication are omitted. Also shown are the multiple correlations (R) representing the net combined effects of the set of background variables for each of the absenteeism measures. The background variables were: (1) Sex, combined with the number of children in the household; (2) Reported problems with the working hours; and (3) Reported problems with transportation to and from work. Three of the four absenteeism measures are shown to be "contaminated" by these background factors.

To assess the residual effects of work role stress upon absenteeism after removal of the variance added (or suppressed) by the background variables, the multiple regression method was again applied, using the four work role stress measures as predictors and with the effects of the background factors controlled. The results appear in Table 9. In sum it can be said that background factors appear to be at least as influential in absenteeism as is work role stress, and that the two sources of influence are not merely duplicative since two significant "predictions" from stress to absenteeism persist after the removal of the effects of background factors. The joint effects of both stress and background variables are represented by

Table 8

Multiple Regressions on Absenteeism Using Background Characteristics as Predictors

Absenteeism	R	B - coefficients		
		Sex and Number of Children	Problems with Hours	Problems with Transportation
Time 1 Reported Absenteeism (N = 627)	.19**	.12**	.13**	-.07*
Time 1 Recorded Absenteeism Month Prior to Interview (N = 519)	.14*	.11**	.08	.01
Time 2 Reported Absenteeism (N = 228)	.12	.00	-.10	-.07
Time 2 Recorded Absenteeism for 1st Six Months (N = 323)	.25**	.14**	.18**	-.11*

*p < .05
**p < .01

Table 9

Multiple Regressions on Absenteeism with Stress Variables as Predictors, Controlling for the Effects of Background Characteristics

		B - coefficients			
Absenteeism	R	Role Ambiguity	Role Overload	Under-Utilization of Skills	Resource Inadequacy
Time 1					
Reported Absenteeism (N = 599)	.14*	.00	.01	.12**	-.07
Time 1					
Recorded Absenteeism Month Prior to Interview (N = 498)	.11	-.11*	.05	.05	-.07
Time 2					
Reported Absenteeism (N = 221)	.15	-.08	-.02	.03	-.14*
Time 2					
Recorded Absenteeism (N = 311 for the first six months)	.21**	.02	.09	.12*	-.11

*p < .05
**p < .01

an R-coefficient of .33 ($p < .001$) compared with .25 for the background variables alone, using six month absenteeism as the criterion.

Similar analyses were conducted using turnover measures as the criterion variables. The results appear in Tables 10 and 11. The pertinent background background variables in this context were: (1) Age; (2) Marital status; (3) Number of children in the household; and (4) Tenure with present employer. Background factors explained a significant part of the variance in three of the four turnover measures, elapsed time to departure being the only criterion that was not significantly implicated. With the removal of the effects of these background factors, work role stress retained a significant relationship to only one of the criteria -- turnover intent at Time 1. The R-coefficient for the turnover intent criterion using both stress and background variables as predictors was $R=.44$, compared with $R=.39$ for background variables alone.

It may be concluded from this segment of the analysis that background factors of the kinds included in this study operate quite strongly to explain withdrawal behavior and appear to duplicate the stress measures in predictive power. The effects of work role stress upon withdrawal behavior are not shown to be pervasive or strong apart from background factors.

Relationship between Absenteeism and Turnover

The primary issue in this phase of the analysis was whether absenteeism and voluntary turnover are to be regarded as alternative forms of withdrawal behavior in which increased episodic absenteeism is a precursor of the ultimate and extended absence called voluntary turnover. The former

Table 10

Multiple Regressions on Turnover Using Background Variables as Predictors

Turnover	R	B - coefficients			
		Age	Marital Status	Number of Children	Tenure
Time 1 Turnover Intent (N = 611)	. 39**	+.19**	-.20**	.02	.12*
Time 2 Turnover Intent (N = 153)	.43**	-.34**	.24**	-.11	.04
Voluntary Turnover (N = 204)	.45**	-.28**	.11	.05	-.17*
Elapsed time to Voluntary Turnover (N = 44)	.34	.33	-.04	.01	-.40*

*p < .05
**p < .01

Table 11

Multiple Regressions on Turnover with Stress Variables as Predictors, Controlling for the Effects of Background Characteristics

Turnover	R	B - coefficients Role Ambiguity	Role Overload	Under-Utilization of Skills	Resource Inadequacy
Time 1 Turnover Intent (N = 585)	.21**	-.06	-.03	-.15**	.09
Time 2 Turnover Intent (N = 148)	.20	.05	-.04	-.10	-.17
Voluntary Turnover (N = 197)	.20	-.01	.08	.13	.14
Elapsed Time to Voluntary Turnover (N = 44)	.27	-.28	.24	-.16	-.03

**p < .01

view implies that the correlation between absenteeism and turnover might characteristically be modest in magnitude or even negative; the latter implies a strong positive correlation and, further, suggests a causal chain in which turnover is best predicted by the combination of generic causes (of withdrawal tendencies) coupled with absenteeism (an indicator that the causes are working).

The correlation between absenteeism and turnover was assessed in two ways. First, the number of absences for the month preceding departure for individuals who voluntarily terminated during the period of the study was compared with the average monthly absenteeism for all others during the first six months of the study period. These figures turned out to be 1.20 (N=50) and .80 (N=289), the difference being statistically significant ($F = 7.49$, $p < .01$). Second, the correlation ratio, allowing for possible curvilinearity, was computed between voluntary turnover and average monthly absenteeism during the first six months of the study period. This correlation ratio (Eta) was .48 (N=217, $p < .01$) indicating a rather strong positive relationship between prior absence rates and subsequent voluntary termination. These results seem unequivocal: increased absenteeism was, for this population, more commonly a precursor to later voluntary turnover than an alternative to turnover.

The hypothesis of a causal chain such that work role stress coupled with early signs of withdrawal (increased absenteeism) would maximally predict later voluntary turnover was given only a limited examination in view of the weak role for stress shown by the preceding analyses. A multiple regression analysis using four work role stresses and actual recorded absenteeism (first six months) as the predictors, and voluntary turnover as the criterion produced a multiple correlation (R) of .34. This added contribution to the prediction by inclusion of the stress measures was very small.

Summary and Interpretations

The initial analyses in this inquiry provided modest but general and significant support for the suggestion that work role stresses have a causal relationship to later withdrawal behaviors of absenteeism and voluntary turnover. This interpretation of the findings is further supported by evidence of a time lag such that stress at time 1 is more closely associated with later absenteeism than with concurrent absenteeism, and by the presence of moderating effects such that the associations are somewhat stronger for individuals with relatively low stress-specific tolerance.

The foregoing results must be evaluated in relationship to the role of demographic and life-contest factors, which are also shown to be associated with withdrawal behavior. Both turnover and absenteeism are statistically predicted about as well from background variables as from the role stress factors. The joint use of both stress and background variables gives only a modest increment in explanatory power over either separately -- more in the case of absenteeism than in the case of turnover. Nevertheless, when the effect of the background variables is removed, there remain significant relationships between stress and both forms of withdrawal. From these results, three interpretations are suggested: (1) The demographic and life-context variables do not work to attenuate the stress-withdrawal relationship, as one might reasonably suppose, but instead they work in concert with the stress factors; (2) The greater force of life circumstances in relation to turnover, as compared with absenteeism, may arise simply because turnover is the more risky and more disruptive of the two forms of withdrawal; (3) There may exist some interactions, not treated here, between a person's background factors and the role stress variables to which the person becomes exposed.

One of the four stresses considered, underutilization of skills, appears to be distinctively different from the others in potency of effect when background factors are partialled out. While the evidence is frail, one can suggest that this source of stress, unlike the others, is unique in that it implies the stressed employee to have marketable skills that diminish the risks that otherwise may accompany absenteeism or voluntary turnover.

The data are explicit on the issue of whether absenteeism generally -- for most of the population although not necessarily all members -- functions as a precursor to voluntary turnover rather than as an alternative to turnover. They correlate positively and significantly. There is evidence that a terminator's absenteeism tends to be elevated compared with the non-terminating population during the period preceding termination.

Only three background factors were treated as significant factors in relation to absenteeism. Female sex combined with number of children in the household is significantly associated with elevated absence rates, as is the experience of problems related to work hours. Transportation problems, however, are shown to have a reversed, although weak, part in absenteeism, perhaps because people with transportation problems get locked into transport arrangements such as car pools that reduce their options for voluntary absence. As to the background factors associated with voluntary turnover, the data show that while people who are older and with long employment tenure more often express an intention to leave, it is the younger and less senior people who tend more to actually undertake voluntary turnover. The older and more tenured may experience both greater risk and lesser opportunity when contemplating turnover.

The foregoing results can be taken as providing some support for the view of March and Simon (1958) which suggests a prepotent role in absenteeism and turnover of opportunity -- i.e., the feasibility of low-risk absenteeism, the cost-benefit balance in absence choices, and the perceived market opportunities for those desiring a change to a better, more convenient job.

The results of this inquiry in general do not provide much encouragement to pursue the issues relating to the function of work role stress in withdrawal behavior unless means are found to enlarge the scope of inquiry and to invoke more complex and potentially more valid models of causation. Future research would be aided by: (1) Consideration of additional forms of work role stress such as, for example, role conflict, degree of promotion and transfer opportunity, and factors that result in the individual being "locked in" to his job; (2) Considering more complex interactions among potential causal variables and interactions that may be applicable to selected subcategories of the employed population; and (3) Improved means for distinguishing truly voluntary from involuntary instances of absence and turnover.

References

Baumgartel, H., & Sobol, R. Background and organizational factors in absenteeism. Personnel Psychology, 1959, 12, 431-443.

Beehr, T. A. Role ambiguity and role stress: Some moderating and intervening variables. Unpublished doctoral dissertation, The University of Michigan, 1974.

Beehr, T. A. Perceived situational moderators of the relationship between subjective role ambiguity and role strain. Journal of Applied Psychology, 1976, 61, 35-40.

Beehr, T. A., & Gupta, N. The structure of employee withdrawal. Organizational Behavior and Human Performance, 1978 (in press).

Behrend, H. Absence and turnover in a changing economic climate. Occupational Psychology, 1953, 27, 69-79.

Boyd, J. B. Interests of engineers related to turnover, selection, and management. Journal of Applied Psychology, 1961, 45, 143-149.

Brayfield, A. H., & Crockett, W. H. Employee attitudes and employee performance. Psychological Bulletin, 1955, 52, 296-424.

Cooper, R., & Payne R. Age and absence: A longitudinal study in three firms. Occupational Psychology, 1965, 39, 31-43.

Evans, W. M. Peer-group interaction and organizational socialization: A study of employee turnover. American Sociological Review, 1963, 28, 436-440.

Farris, G. F. A predictive study of turnover. Personnel Psychology, 1971, 24, 311-328.

Guest, R. H. A neglected factor in labor turnover. Occupational Psychology, 1955, 29, 217-231.

Hackman, J. R., & Lawler, E. E. III. Employee reactions to job characteristics. Journal of Applied Psychology, 1971, 55, 259-286.

Herzberg, F., Mausner, B., Peterson, R. D., & Capwell, D. F. Job attitudes: Review of research and opinion. Pittsburgh: Psychological Service of Pittsburgh, 1957.

Hewitt, D., & Parfitt, J. A note on working morale and size of group. Occupational Psychology, 1953, 27, 38-42.

Hill, J. M., & Trist, E. L. Changes in accidents and other absences with length of service. Human Relations, 1955, 8, 121-152.

Hulin, C. L. Effects of changes in job-satisfaction levels in employee turnover. Journal of Applied Psychology, 1968, 52, 122-126.

Huse, E. F., & Taylor, E. K. The reliability of absence measures. *Journal of Applied Psychology*, 1962, *46*, 159-160.

Ingham, G. Size of industrial organization and worker behaviour. Cambridge: Cambridge University Press, 1970.

Katzell, M. E. Expectations and dropouts in schools of nursing. *Journal of Applied Psychology*, 1968, *52*, 154-157.

Kerr, W. A., Koppelmier, G. J., & Sullivan, J. J. Absenteeism, turnover, and morale in a metals fabrication factory. *Occupational Psychology*, 1951, *25*, 50-55.

Kim, S., Roderick, R., & Shea, J. Dual careers: A longitudinal study of the labor market experience of women. Volume 2. Washington, D.C.: United States Government Printing Office, 1973.

Lawler, E. E. III., & Hackman, J. R. Impact on employee participation in the development of pay incentive plans: A field experiment. *Journal of Applied Psychology*, 1969, *53*, 467-471.

Ley, R. Labour turnover as a function of worker differences, work environment, and authoritarianism of foreman. *Journal of Applied Psychology*, 1966, *50*, 497-500.

Lyons, T. F. Role clarity, need for clarity, satisfaction, tension and withdrawal. *Organizational Behavior and Human Performance*, 1971, *6*, 99-110.

Lyons, T. F. Turnover and absenteeism: A review of relationships and shared correlates. *Personnel Psychology*, 1972, *25*, 271-281.

McNemar, Q. *Psychological Statistics*. New York: Wiley, 1969.

Macy, B. A., & Mirvis, P. H. A methodology for assessment of quality of work life and organizational effectiveness in behavioral-economic terms. *Administrative Science Quarterly*, 1976, *21*, 212-226.

Mangione, T. W. Turnover: A model and review of the literature. Ann Arbor: Survey Research Center, 1972. (Mimeo)

March, J. G., & Simon, H. A. *Organizations*. New York: Wiley, 1958.

Melbin, M. Organizational practice and individual behavior: Absenteeism among psychiatric aides. *American Sociological Review*, 1961, *26*, 14-23.

Metzner, H., & Mann, F. C. Employee attitudes and absences. *Personnel Psychology*, 1953, *6*, 467-485.

Parnes, H., Zeller, F., Shea, J., Kohn, A., & Meyer, J. Career thresholds: A longitudinal study of the educational and labor market experience of male youth. Vol. 2. Washington, D.C.: United States Government Printing Office, 1971.

Porter, L. W., & Steers, R. M. Organizational, work, and personal factors in employee turnover and absenteeism. Journal of Applied Psychology, 1973, 80, 151-176.

Price, J. L. Handbook of organizational measurement. Lexington: Heath, 1972.

Quinn, R. P., Fine, B. D., & Levitin, T. E. Turnover and training: A social psychological study of disadvantaged workers. Springfield: National Technical Information, Operations Division, Document #:PB194775, 1970.

Quinn, R. P., & Shepard, L. J. The 1972-73 quality of employment survey: Descriptive statistics, with comparison data from the 1969-70 survey of working conditions. Ann Arbor, Mi.: Survey Research Center, 1974.

Robinson, D. D. Prediction of clerical turnover in banks by means of a weighted application blank. Journal of Applied Psychology, 1972, 56, 282.

Roderick, R., & Davis, J. Years of decision: A longitudinal study of the educational and labor market experience of young women. Columbus: The Ohio State University, 1973.

Saleh, S. D., Lee, R. J., & Prien, E. P. Why nurses leave their jobs -- An analysis of female turnover. Personnel Administration, 1965, 28, 25-28.

Schefler, K. C., Lawler, E. E. III, & Hackman, J. R. Long-term impact of employee participation in the development of pay incentive plans: A field experiment revisited. Journal of Applied Psychology, 1971, 55, 182-186.

Schuh, A. The predictability of employee tenure: A review of the literature. Personnel Psychology, 1967, 20, 133-152.

Smith, P. C. Behaviors, results, and organizational effectiveness: The problem of criteria. In M. D. Dunnette, (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

Turner, A. N., & Lawrence, P. R. Industrial jobs and the worker: An investigation of response to task attributes. Boston: Harvard University Press, Division of Research, 1965.

Vroom, V. Work and motivation. New York: Wiley, 1964.

Weitz, J. Job expectancy and survival. Journal of Applied Psychology, 1956, 40, 245-247.

White, B. L. Job attitudes, absence from work, and labor turnover. Personnel Practices Bulletin, 1960, 16, 18-23.

Chapter 16

EFFECTIVENESS OF WITHDRAWAL STRATEGIES IN COPING WITH WORK ROLE STRESS

by

Terry A. Beehr

ABSTRACT

This inquiry aimed to explore the extent to which withdrawal from the work situation, either psychological withdrawal or physical withdrawal, might serve as a means for moderating the personal strains arising from work role stresses. The analysis focused upon individuals known to be in stressful jobs, and examined the extent to which various forms of withdrawal are associated with evidence of reduced individual strain.

Stressful jobs were defined by high scores on role ambiguity, role overload, inadequacy of work resources, and underutilization of skills. Six measures of strain were considered: depressed mood, work-related illness or injury, poor physical health, low self-esteem, life dissatisfaction, and job dissatisfaction. Withdrawal was defined to include absence, lateness, changing jobs, and psychological withdrawal in the form of low rates of interaction with supervisor and fellow workers.

While the results of this exploratory inquiry must be interpreted with caution, it appears that stressed individuals who are experiencing strain do indeed often withdraw in one or several of the above-mentioned ways. There is no evidence, however, that these withdrawal strategies actually reduce strain. While some strain reduction did occur over time, it appears to arise from other sources, not from withdrawal. It is suggested that withdrawal may be not only ineffective, but possibly self-defeating as a coping strategy.

Chapter 16

EFFECTIVENESS OF WITHDRAWAL STRATEGIES IN COPING WITH WORK ROLE STRESS

A decade of research on organizational and job stress has amply demonstrated that it is a worthy topic, both for basic inquiry into issues of human tolerance and adaptivity, and also for more immediately practical inquiry relating to the optimizing of organizational and job environments. Work role stresses are known to have implications for role performance and for the job occupant's personal welfare; both employers and their employees, not to mention society at large, have a stake in the matter. Job stresses are known to be associated with adverse physical and mental states of persons (strain), and with outcomes of impaired mental health, poor work performance, and the rate and severity of some illnesses and injuries (Caplan, et al., 1975; French & Caplan, 1973; Kasl, 1973; Kornhauser, 1965; Kahn, et al., 1964).

Many of the prevalent and significant work role stresses have been identified and partially assessed. Those taken into account in the principal review (Kasl, 1973) and the most inclusive empirical inquiry (Caplan, et al., 1975, pages 36-38) can be summarized briefly as follows:

1. Working conditions, such as exposure to health or safety hazards, unpleasant working conditions, and the like.

2. Demands of the work itself, such as overload, excessive pace or physical effort, excessive responsibilities, concentration, repetitious or boring work, underutilization of skills, inadequate resources for doing the job, and others.
3. Hours of work, including variable shifts, excessive hours, non-standard hours that conflict with family and other off-job activities, etc.
4. Organizational factors, such as ambiguous or conflicting role demands, excessively close or insufficient supervision, lack of information relating to the job and one's performance, insufficient or excessive autonomy, insufficient influence upon matters relating to the work.
5. Rewards, such as insufficiencies or inequities in pay, inadequate promotional opportunities, poor career development prospects.

Employees exposed to stressful conditions are not necessarily passive victims inevitably experiencing strain. A common response to stress is to cope with it, one way or another, to the end of minimizing or avoiding the stress, or moderating its adverse strain effects. In the individual's choice of coping strategies some conflict of interest may appear between the worker and the employer, for some coping strategies may have the effect of protecting and preserving the individual's capacities for immediate work role fulfillment while others directly or indirectly constrain role fulfillment.

Among the potential means for coping with work role stress are a set that have in common the separation or distancing of the worker from his work location and the people in it. Withdrawal, psychological or physical, is the common theme. This report will be concerned with seven forms of withdrawal from the job: turnover (voluntarily quitting the employer entirely), absence, lateness, reduced job involvement, reduced interaction with supervisor, reduced interaction with co-workers, and changing jobs but remaining with the same employer. All have an element of voluntary action or attitude formation by the worker, and all carry an implication of reduced amount of quality of work role fulfillment.

The purpose of the inquiry was to examine a population of employees known to be exposed to relatively high amounts of work role stress to determine whether those engaging in the withdrawal behaviors had been more successful than others in moderating strain outcomes. The model upon which this analysis is based states that in the presence of relatively high stress, an employee will experience relatively high strain unless that strain is moderated by successful coping; coping might be aided, among other ways, through partial or total withdrawal from the work role and work environment. Four types of job stress were considered: role ambiguity, role overload, underutilization of skills, and inadequate resources for doing the job. These have been identified previously as determinants of job-related strains (e.g., Margolis, Kroes, & Quinn, 1974; French & Caplan, 1973). Six measures of strain were used: physical health (adjusted for sex and age), work-related illness and injury, depressed mood, low self-esteem, job dissatisfaction, and dissatisfaction with life.

Method

Sample

The sample comprised 651 respondents employed 35 hours or more per week by five midwestern work organizations: a printing company (n = 173), a small research and development company (n = 21), two automotive supply companies (Ns = 120 and 124), and four services departments of a hospital (N = 213). These organizations were chosen in part on the criterion that they contain a diverse set of jobs, and there is evidence that the occupational census codes of the respondents roughly match the codes of a national sample of workers (Beehr, 1974). All supervisors were included in the sample, and nonsupervisors were sampled systematically at rates varying from organization to organization between 25% and 100%. The response rate was 72.9%.

Fifty-one percent of the respondents were male, 79.8% were white, 68.5% were married, 73.0% had finished at least high school, and the mean age was 35 years. Comparable information on the nonrespondents was not available.

Data Collection

Data were collected at two times and in four ways. A main source of data was a 90-minute structured interview in the respondent's home conducted during the winter of 1972-1973. Concurrent Time 1 data were obtained from trained observers who visited the workplaces on a prescheduled basis and reported their observations on a standardized report form. Supervisory ratings of the performance of subordinates included a report on lateness to work, but these ratings were not available for all segments of the population. In addition, employers' records provided information on absences. A second wave of measurements was conducted about 20 months later. These Time 2 measures included a repetition of the observation and interview procedures, and a record search for information about absences, internal job changes and voluntary departures.

Tables 1 and 2 list the variables employed in this inquiry, indicating the data source for each, as well as its mean, standard deviation and scale range. The number of cases varies for the several measures, and is substantially smaller for the Time 2 data reported in Table 2. The measures are described below in the order of their first appearance in these tables. Intercorrelations among the strain measures (Time 1) are shown in Table 3.

Stress Measures - Interview

Role ambiguity was measured by a four-item index composed of responses to the following questions. Responses were given on a four-category scale ranging from "Very true" to "Not at all true". All items but the third listed were taken from the Quality of Employment Survey (Quinn & Shepard, 1974). Index reliability (Alpha) was .69.

How true is this of your job?

I can predict what others will expect of me tomorrow.

I am clear on what others expect of me on my job.

On my job, whatever situation arises, there are procedures for handling it.

I get enough facts and information to work my best.

Role overload was measured by combining responses to two questions on a four-category scale ranging from "Very true" to "Not at all true". Respondents were asked how true each of the following was of their own jobs: "I am given enough time to do what others expect of me."; "I am able to

Table 1

Mean, Standard Deviation, and Range for All Variables (Time 1)

Variable	N	Mean	Standard Deviation	Possible Range
Stresses				
Interviews				
Role ambiguity	650	1.92	0.64	1-4
Role overload	641	1.79	0.68	1-4
Resource inadequacy	637	1.73	0.57	1-5
Underutilization of skills	642	1.35	1.87	-7 - +7
Observations				
Role ambiguity	579	3.14	1.01	1-7
Role overload	579	2.47	0.55	1-7
Resource inadequacy	579	2.49	0.48	1-6
Underutilization of skills	579	4.44	1.54	1-7
Strains				
Depressed mood	648	1.85	0.47	1-7
Work-related illness and injury	645	1.40	1.00	1-5
Poor overall physical health	631	3.51	1.41	1-5
Low self-esteem	647	1.98	1.05	1-7
Life dissatisfaction	646	0.14	0.91	-3.6 - +1.4
Job dissatisfaction	650	2.45	1.03	1-5
Withdrawal				
Interview				
Absenteeism	641	1.55	0.73	1-5
Lateness	622	1.57	6.70	1-5
Low Involvement	647	2.94	0.92	1-5
Low interaction with supervisor	628	1.20	0.89	1-4
Low interaction with coworkers	649	2.40	1.30	0-5
Records				
Absenteeism	531	1.09	1.45	0-9
Supervisor ratings				
Lateness	286	1.98	1.29	1-7

Table 2

Mean, Standard Deviations, and Range for All Variables
(Time 2)

Variable	N	Mean	Standard Deviation	Possible Range
Stresses				
Interviews				
Role ambiguity	261	1.88	0.60	1-4
Role overload	259	1.63	0.64	1-4
Resource inadequacy	262	1.81	0.56	1-4
Underutilization of skills	264	1.41	1.92	-7 - +7
Observations				
Role ambiguity	100	3.16	1.45	1-7
Resource inadequacy	100	1.89	0.74	1-4
Underutilization of skills	100	4.29	1.85	1-7
Strains				
Depressed mood	269	1.84	0.47	1-7
Work-related illness and injury	267	1.39	1.01	1-5
Poor overall physical health	268	3.29	1.45	1-5
Low self-esteem	268	1.96	0.99	1-7
Life dissatisfaction	268	-1.36	0.59	-2.6 - 0
Job dissatisfaction	268	2.38	0.94	1-5
Withdrawal				
Records				
Voluntary turnover	286	1.17	0.38	1-2
Intraorganizational job change	236	1.31	0.46	1-2

Table 3

Intercorrelations among Strain Measures
(Time 1, N=644)

	1	2	3	4	5	6
Depressed mood	(.71)					
Illness/injury	.09*	--				
Physical health	.22**	.17**	--			
Low self-esteem	.31**	.01	.02	(.68)[a]		
Life dissatisfaction	.39**	.09*	.16**	.28**	(.79)[a]	
Job dissatisfaction	.43**	.13**	.15**	.25**	.27**	(.80)[a]

* p < .05

** p < .01

[a] Figures in parentheses are internal consistency estimates (alpha), omitted where not available.

complete the work I start." This index had very low reliability compared to others employed, namely, .47 but was judged suitable for use. The items were taken from Quinn & Shepard (1974).

Resource inadequacy was measured by an index comprising four items known by factor analysis to be associated within a common "resource" dimension (Quinn & Shepard, 1974). Responses were on four-category scales ranging from "Very true" to "Not at all true" for the first three items, and from "Great, couldn't be better" to "Not very good" for the fourth item. The questions were:

> (My supervisor) knows his/her job well.
>
> (My supervisor) maintains high standards of performance in his/her own work.
>
> I am given enough machinery and equipment to do my best.
>
> How good are they (your work group) about giving you the help you need to do your job well?

Underutilization of skills was measured by an index derived from three interview questions:

> What is the level of school or college you feel is needed by a person in your job?
>
> What was the highest grade of school or year of college you have completed?
>
> Through your previous experience and training, do you have some skills that you would like to be using in your work but can't use on your present job?

The first two, answered on common eight-category scales although at different times in the interview, were used to derive a discrepancy score, and this was combined with a scaled response to the third question to form the index.

Stress Measures - Observation

The stress measures obtained by observation were based upon the reports of trained observers, not otherwise associated with the workplace, who visited the subject's work location on a prescheduled basis for periods of observation. A structured questionnaire-report provided a number of rating scales on which the observers recorded their judgements. The rating scales employed to report work role stress were all six-category or seven-category scales with descriptive labels representing frequencies, amounts, intensities, etc. The report on pressures for better performance, for example, included the rating category #7: "The individual is being constantly exposed to demands for working harder or doing a better job". A simpler rating scale, e.g., that responsive to a statement concerning the availability of tools, machinery and equipment, ranged from "Very true" to "Not at all true". The scale relating to degree of uncertainty in the work role, for example, included rating category #1, "Very little; the individual almost always knows what to expect and is never surprised by something happening unexpectedly on the job". A full description of the observer training course, the reporting forms, and the reliability observations is provided in Chapter 4. The items comprising each of the observation stress measures were:

Role ambiguity

How much uncertainty is there in the job?

The job requires the individual to be prepared to handle surprising or unpredictable situations.

The job is one that is highly predictable, and that rarely presents the individual with surprising or unexpected problems.

The individual working on the job does tasks which are clearly defined.

Role overload

Are there any pressures for better performance over and above what is reasonable?

The individual doing the job is asked to do excessive amounts of work.

He/she has enough time to do what he/she is expected to do.

Resource inadequacy

He/she is given enough space to do his/her job.

He/she is given adequate lighting for his/her particular job.

He/she has adequate access to machinery, tools and other equipment.

The individual working on the job frequently has to stop to get things he/she needs and doesn't have readily available.

Underutilization of skills

To what extent does the job require the use of sophisticated or complex skills?

How intellectually demanding is the job?

The job is so simple that virtually anybody could handle it with little or no initial training.

Strain Measures - Interview

All six of the role strain measures were derived from the prior scale development and validation work reported by Quinn & Shepard (1974). In some cases, as noted below, a subset of the original index components was used, and two were reversed in sign for consistency in this report.

Depressed mood was measured by a ten-item index. The question form and components are shown below. Response categories ranged from "Often" to "Never".

How often do you feel this way at work?

I feel downhearted and blue.

I get tired for no reason.

I find myself restless and can't keep still.

I find it easy to do the things I used to do.

My mind is as clear as it used to be.

I feel hopeful about the future.

I find it easy to make decisions.

I am more irritable than usual.

I still enjoy the things I used to.

I feel that I am useful and needed.

Work related illness and injury was measured by combining responses to two questions: "Within the last three years have you had any illnesses or injuries you think were caused or made worse or more severe by any job you had during this period? Could you tell me what those illnesses and injuries were?" "When you had (this illness or injury) did it keep you away from your job for more than two weeks?" The index was composed of the number of illnesses and injuries reported to the first question, corrected for tenure on the present job, plus a scale step increment if a two-week absence had occurred on his present job.

Overall poor physical health was measured by the coding of responses to a rather long series of questions dealing with (1) recent symptoms of ill health, (2) treatment for health conditions during the last year, and (3) self report of sense of health and energy level. The index code categories range from (5) Having been under treatment during the last year for two or more chronic conditions, to (1) Reporting no chronic conditions nor symptoms and also reporting good health and high energy. This index was based upon the work of Belloc, et al. (1971), modified to suit application to an employed population.

Low self-esteem was measured by a three-item subset from the Quinn & Shepard (1974) semantic differential scale. The anchors on seven-point scales were "successful-unsuccessful", "important-unimportant", and "doing my best-not doing my best", in response to the lead-in question, "How do you see yourself in your work?". The index has a reliability (alpha) of .68.

The life dissatisfaction measure consisted of a nine-item subset of Quinn & Shepard's (1974) life satisfaction index. Seven of these items were answered on a semantic differential scale. The seven pairs of anchors were "interesting-boring", "enjoyable-miserable", "friendly-lonely", "full-empty", "hopeful-discouraging", "rewarding-disappointing", and "brings out the best in me-doesn't give me much of a chance" in response to the question, "How do you feel about your present life in general?". The fixed-alternative questions were:

> In general, how satisfying do you find the ways you're spending your life these days? Would you call it completely satisfying, pretty satisfying, or not very satisfying?

> Taking all things together, how would you say things are these days? Would you say you're very happy, pretty happy, or not too happy these days?

The life dissatisfaction index had a possible range of 1-7, and a reliability of .79.

Job dissatisfaction was measured by a four-item subset of Quinn & Shepard's (1974) global job satisfaction index:

> All in all, how satisfied would you say you are with your job -- very satisfied, somewhat satisfied, not too satisfied, or not at all satisfied?
>
> In general, how well would you say that your job measures up to the sort of job you wanted when you took it? Would you say it is very much like, somewhat like, or not at all like the job you wanted when you took it?
>
> Knowing what you know now, if you had to decide all over again whether to take the job you now have, what would you decide? Would you decide without any hesitation to take the same job, would you have some second thoughts, or would you decide definitely not to take the same job?
>
> If a good friend of yours told you that he/she was interested in working in a job like yours, what would you tell him/her? Would you strongly recommend this job, would you have doubts about recommending it, or would you strongly advise him/her against this sort of job?

The job dissatisfaction index had a possible range of 1-5, and a reliability of .80.

Withdrawal Measures - Interview

Absence was measured in the interview by an index comprising three questions:

> Would you say you are absent from work more often than other people you work with, less often, or about the same?
>
> Aside from any paid vacation and holidays, how many days of scheduled work have you missed in the past month?
>
> How many of these days did you miss just because you didn't feel like going to work that day?

Lateness was measured in the interview with an index comprising two questions, namely the last two shown immediately above, but with reference to lateness during the last two weeks.

Job involvement was measured in the interview with a single question: "Some people are completely involved in their job -- they are absorbed in it night and day. For other people, their job is simply one of several interests. How involved do you feel in your job -- very little, slightly, moderately, or strongly involved?".

Interaction and communication with supervisor was measured by combining responses to two statements, with responses on four-point scales ranging from "Very true" to "Not at all true". The items were: "My supervisor goes out of his/her way to praise good work", and "My supervisor lets his/her subordinates know how well they are doing their jobs".

Interaction and communication with co-workers was measured in the interview with three questions:

> Is there any group of people at (your place of work) that you think of as your co-workers -- people whom you see just about every day and with whom you have to work closely in order to do the job well?
>
> How well do you feel they help each other on the job?
>
> How good are they about giving you the help you need to do your job well?

Those responding "No" to the first question were assigned a code of 1, while others received their average score on the succeeding questions, each having response categories ranging from "Not very good" to "Great, couldn't be better".

Withdrawal Measures - Non-Interview

Absence was measured by an actual count from the employers' records of the number of occasions during the past month in which the employee was absent from scheduled work for a period of a day or more.

Lateness was measured by including reference to lateness behavior in a semantic differential question incorporated in an employee rating form. This form was filled out by the immediate supervisor for each of his/her subordinates. This item of information is missing for some of the population owing to failure to gain consent to the procedure from all employers, and to instances where the supervisor or subordinate had been in their relationship too short a time to allow rating. The polar anchors on the seven-point semantic differential item were "Always arrives on time" and "Always late".

Voluntary turnover and intraorganizational job change were measured at Time 2 (only) by a search of the employers' records. For the turnover variable, individuals were coded 1 if still on the original Time 1 job and 2 if they had left to join another employer. The other variable was coded similarly.

Analysis Plans

Since little was known in advance about the measures employed and about their interrelationships, two critical decisions were made at the start: (1) The several variables representing each of the three main classes (role stresses, role strains and forms of withdrawal) would be treated separately and not be combined into more inclusive indices, and (2) The analyses would focus upon the relatively highly stressed and strained segment of the population by population subsetting rather than treating the population as a whole. These choices have certain advantages in an exploratory inquiry as well as important disadvantages; some reflections on them appear later in this report.

Three analyses were performed, the first using only Time 1 data, and the other two using as well the data from Time 2 measurements.

Analysis 1 aimed to compare the role strain outcomes of people who shared the condition of high role stress but differed in their use of each of the withdrawal modes. The operations were, in sequence, to segregate the subjects above the population median on a given stress variable, to divide these at the median of a given withdrawal measure, to compute the mean scores on each strain measure for these two compared groups, and to determine the difference between means and its significance. With eight stress measures, seven withdrawal measures, and six strain measures, a total of 336 such analyses were made.

Analysis 2 focused upon those subjects who were exposed to high stress and experienced high strain at Time 1 and who remained in the same job at least until Time 2. Such people would be likely during subsequent months to be using withdrawal strategies as a means to reduce strain, with the effect of some strain reduction by Time 2. The operations were, in sequence, to segregate those above the population median on both a given stress and a given strain at Time 1 and also above the population median on the stress variable at Time 2 (i.e., role stress remained high), to subset these subjects according to their level of strain at Time 2, and to compare those with high vs. low strain at Time 2 with respect to their use at Time 1 of the several withdrawal modes. Whereas Analysis 1 sought the concurrent effects of withdrawal, Analysis 2 sought delayed effects. The number of tests was potentially 294.

Analysis 3 focused upon those subjects who had changed employers or jobs or both between the Time 1 and Time 2 measurements. The guiding supposition was that subjects under role stress or role strain at Time 1 would be more likely than others to change jobs as a means for stress avoidance and strain reduction. Those who left their jobs and employers voluntarily should accordingly have been at Time 1 among those subjected to relatively high role stress and experiencing high role strain. Those who changed jobs with the same employer should display reduced stress and/or strain at Time 2 to the extent (1) that their Time 1 stress and strain was relatively high, and (2) that the change was successful as a coping strategy. It was not possible to test the effectiveness of turnover (changing jobs and employers) as a coping strategy, as the departed persons were not accessible for Time 2 measurement.

Significance tests of mean differences were uniformly based upon two-way analyses of variance. Significance values are expressed as probabilities with the conventional $p < .05$ taken as a standard for judging a difference to be significant.

Some Issues in Analysis

The conception of social and psychological processes that underlie the foregoing analysis strategies holds that stress causes strain, and that strain induces coping behavior (e.g., withdrawal) which if successful moderates the strain or the stress or both. Two difficulties are posed by such a conception when investigation is through non-experimental methods.

Time is an obvious problem, for the model is not one of discrete observable events, but rather a model of continuous change. In the present context, the rub occurs at two places: (1) The time span between Time 1 and Time 2 measures is an arbitrary one, fixed for reasons not related to the present analysis, and it may be an unsuitable time period; (2) The proposed causal direction of relationship between role strain and use of withdrawal as a means for coping is arbitrary, as is the case with any feedback model in which there may occur an alternation or mutuality of causation. Since we cannot know what direction of causation is caught by the timing of our concurrent data, we can with certainty show if a relationship exists, but we cannot test bi-directionality of causation.

The second issue arises because we consider only one category of coping behavior -- withdrawal in various forms -- and have no way of knowing whether other forms of coping are complementary, or instead are mutually exclusive alternatives. Thus, it must be assumed that some attenuation of relationships may be present arising from the existence of coping strategies not treated in this analysis.

Results

Analysis 1

The extensive results from Analysis 1 are not presented in full detail, but are summarized in Table 4. Of the 336 tests of association between role strain and use of withdrawal options (high stress subjects only), 152 -- 45% of them -- are significant at the five per cent level of probability. There is surely a pervasive relationship between levels of strain and rates (or degrees) of usage of withdrawal strategies by people in stressful jobs.

An inspection of the source tables shows that the direction of the differences tested is consistently that under high stress conditions high strain is concurrently associated with high withdrawal, not low. Only one of the significant differences reverses that direction, even though in our model reversals are equally possible and plausible. It appears from these concurrent measures that withdrawal is not shown to be effective in moderating strains associated with role stress.

There are few variations in the strength of these relationships as between the different stresses, strains, and forms of withdrawal (Table 3). Relatively high proportions of significant differences occur in connection with one strain, depressed mood (68%), and three withdrawal measures, absence (65%), job involvement (65%), and low interaction with supervisor

Table 4

Summary Table for Analysis Strategy One: Significant F's for Relationship between Withdrawal Techniques and Strains for Employees Experiencing High Role Stress

Strains, Stresses and Withdrawal Indicators	Number of F's	Percent of F's p <.05
Strains		
Depressed mood	56	67.9
Work-related illness and injury	56	35.7
Poor overall physical health	56	32.1
Low self-esteem	56	39.3
Life dissatisfaction	56	37.5
Job dissatisfaction	56	58.9
Total	336	45.2
Stresses		
Interview		
Role ambiguity	42	57.1
Role overload	42	40.5
Resource inadequacy	42	47.6
Underutilization of skills	42	42.9
Subtotal	168	47.0
Observation		
Role ambiguity	42	40.5
Role overload	42	33.3
Resource inadequacy	42	40.5
Underutilization of skills	42	59.5
Subtotal	168	43.5
Total	336	45.2
Withdrawal		
Interview		
Absenteeism	48	64.6
Lateness	48	2.1
Involvement	48	64.6
Low interaction with supervisor	48	81.3
Low interaction with coworkers	48	25.0
Subtotal	240	47.5
Records		
Absenteeism	48	54.2
Supervisor ratings		
Lateness	48	25.0
Total	336	45.2

$p<.05$

(81%). Relatively low frequencies of significant relationships are associated with three withdrawal measures: Lateness (two measures, 2% and 25%) and interaction with co-workers (25%). Some of these variations may be connected with the differences in measurement reliability, although this association cannot be tested because some of the measures lack reliability estimates. The two main sources of data -- interviews and observers' ratings -- do not appear to be systematically different in their capacity to deliver significant differences.

Analysis 2

Analysis 2 proved to be ineffective in one important sense because the population, the jobs, and the time period, failed to provide the variations required by the analysis plan. The plan required a sufficient number of cases with high stress throughout the period of study (20 months), and also sufficient diversity in role strain at Time 2, to allow comparisons to see if withdrawal strategies at Time 1 could explain strain reduction at Time 2. However, of the 294 potential tests, 112 provided no low-strain cases at all at Time 2, and only five provided enough cases in each of the two required classifications to permit significance testing. Many of "the same" jobs, initially stressful, had become low-stress jobs by Time 2 according to both interview and observer data.

Nevertheless, this outcome of the analysis is not lacking in meaning, for of the workers initially under conditions of high role stress and high role strain and who continued to be under high role stress, very few by Time 2 had found any effective means (whether by withdrawal or other means) for moderating their role strain. So few reported strain reduction that it proved impossible to assess the impact, if any, of withdrawal coping strategies.

Analysis 3

Table 5 compares employees staying on the same job with those promoted or transferred within the same firm as to their stress and strain levels at Time 2, treating only those who had relatively high stress and strain at Time 1. Of the eleven tests, only one is significant. That one is readily understandable, as it occurs in connection with the role stress of underutilization of skills as measured in the interviews. One may presume that most of these internal job changes were voluntary shifts to more demanding jobs which reduced the degree of skill underutilization experienced by the worker. The table as a whole shows that intraorganizational job changing is not otherwise an effective means for reducing role stresses or reducing role strains in this population.

Table 6, which compares those workers who left their employers with those who stayed on in their original Time 1 job, shows only slight evidence that this type of job change is associated significantly with Time 1 levels of role stress. The one significant association relates to the role stress of underutilization of skills (interview), a result compatible with the speculation offered above that many such job changes are from less to more demanding jobs. However, as to role strain the results are different, for three strain indicators out of five (job dissatisfaction, life dissatisfaction, and depressed mood) show that the leavers were, in advance of leaving, experiencing more strain than their co-workers who stayed on.

Discussion

We have attempted to explore a recursive model in which role stress is

Table 5

Time 2 Stresses and Strains for Employees with Strong (above Median) Corresponding Stresses and Strains in Time 1: Comparison of Employees on Same Job in Same Company with Employees on Different Job in Same Company

Employees	Stress or Strain Measure on which Employees Were High in Time 1	Time 2 Stress or Strain		n	F	Significance of Difference (p)
		Mean	Standard Deviation			
Same job	Interview Role Ambiguity	2.04	0.58	75	0.48	.49
Different job		2.13	0.62	34		
Same job	Interview Role Overload	1.81	0.69	59	0.00	.95
Different job		1.80	0.58	32		
Same job	Interview Underutilization of Skills	1.98	1.88	58	3.92	.05
Different job		2.88	2.12	28		
Same job	Interview Resource Inadequacy	1.97	0.51	78	0.87	.35
Different job		1.87	0.58	36		
Same job	Observation Resource Inadequacy	1.79	0.69	47	2.06	.16
Different job		2.14	0.77	10		
Same job	Work-related Illness and Injury	1.77	1.39	26	0.87	.36
Different job		2.21	1.78	19		
Same job	Depressed Mood	1.98	0.51	84	0.09	.77
Different job		2.01	0.47	40		
Same job	Life Dissatisfaction	1.20	0.22	74	1.27	.26
Different job		1.09	0.15	37		
Same job	Low Self-esteem	2.17	0.72	62	0.19	.67
Different job		2.25	0.98	34		
Same job	Job Dissatisfaction	2.89	0.83	63	0.45	.51
Different job		2.77	0.94	37		
Same job	Poor Overall Physical Health	3.72	1.32	90	0.58	.45
Different job		3.53	1.46	45		

Table 6

Time 1 Stresses and Strains for Employees on Same Job in Same Company in Time 2 with Employees Who Turned Over Voluntarily

Employees	Time 1 Stress or Strain	Time 1 Stress or Strain Mean	Time 1 Stress or Strain Standard Deviation	n	F	Significance of Difference (p)
Same job	Interview Role Ambiguity	1.89	0.61	235	0.25	.62
Turned over		1.85	0.63	50		
Same job	Interview	1.69	0.64	232	1.67	.20
Turned over	Role Overload	1.82	0.66	50		
Same job	Interview	1.19	1.80	235	15.07	.00
Turned over	Underutilization of Skills	2.30	1.98	50		
Same job	Interview Resource Inadequacy	1.77	0.57	229	0.45	.50
Turned over		1.67	0.48	50		
Same job	Observation	3.31	1.10	220	0.01	.93
Turned over	Role Ambiguity	3.29	1.14	43		
Same job	Observation	2.58	0.50	220	0.12	.73
Turned over	Resource Inadequacy	2.60	0.59	43		
Same job	Observation	4.46	1.57	220	0.48	.49
Turned over	Underutilization of Skills	4.27	1.87	43		
Same job	Poor Overall Physical Health	3.42	1.48	230	0.78	.38
Turned over		3.62	1.41	50		
Same job	Work-related Illness and Injury	1.42	1.02	236	0.22	.64
Turned over		1.50	1.17	50		

(continued..................)

Table 6--(continued)

Employees	Time 1 Stress or Strain	Time 1 Stress or Strain Mean	 Standard Deviation	n	F	Significance of Difference (p)
Same job	Job Dissatisfaction	2.28	0.96	235	10.44	.00
Turned over		2.78	1.06	50		
Same job	Life Dissatisfaction	-0.03	0.85	232	8.16	.00
Turned over		0.35	0.91	50		
Same job	Depressed mood	1.83	0.44	233	3.95	.00
Turned over		1.97	0.46	50		

seen as causing role strain, which in turn induces coping strategies which reduce the strain. The testing of the model is incomplete in significant ways, but certain aspects of it have been illuminated.

The first analytic method provided clear results compatible with the linkage of role strain to adoption of withdrawal coping strategies. Since the analysis was concurrent, the suggested direction of causation is not proven. The association is not an artifact of shared measurement error, for it appeared as well under conditions of independent measurement.

The second analytic method provided too few cases to allow valid statistical analysis as intended: it could not be shown that Time 1 withdrawals did or did not achieve a reduction in strains by Time 2. The second analysis did show that in those jobs that did remain stressful over a span of time there occurred little or no diminution of role strain for the occupants, suggesting that the concurrent linkage of role stress to strain is a persistent one not commonly moderated by the passage of time nor by the various coping strategies that the occupants may have attempted to use. If the job remains stressful the occupant remains strained.

The effectiveness of voluntary turnover as a coping strategy could not be assessed in this study, but it is confirmed, as expected, that those workers choosing to leave had experienced more role strain than those who chose to stay on, even though the initial differences as to role stress were insignificant except as to perceived skill underutilization. It appears that those individuals with an unusually strong strain response to stress were prone to leave. As to internal job changes, there is no evidence in this study that this is an effective coping strategy except with respect to the stress of skill underutilization. However, one must consider that some gains for the intraorganizational job changers may have occurred that are offset by the temporarily added stresses and strains incidental to the process of job change itself.

Two lines of further inquiry are suggested by the findings of this study in conjunction with the limitations of its data.

If strategies exist that allow successful coping with work role stresses and strains by the worker himself, it is likely that they will prove to be strategies other than the negative and possibly self-defeating ones of withdrawal. Other coping strategies need to be investigated. One thinks of the strategies that are the opposite, in the sense of reflecting intensified engagement with the job rather than distancing from it. Examples of such coping strategies are: (1) Engagement in processes of workplace participation and control-sharing that offer the possibility of directly altering the stressful properties of work roles; (2) Mutual social support strategies that appear from fragmentary evidence to be directly associated with strain reduction (see Chapter 5).

The results of this inquiry, and particularly the changing levels of both stress and strain over time in a set of jobs continuously occupied by the same people, suggests that the stress and strain may be linked with job tenure in such a way that any newly-taken job is stressful and strain-generating while any job continuously occupied tends to become less so or to stimulate self-selected instances of turnover. Such a phenomenon might account for the short run absence of coping success in connection with intra-organizational job changes. It suggests that there may be a high cost to workers in making job changes that do not provide other advantages to offset the stress and strain costs.

Some Methodological Comments

The analyses reported here do not exhaust the potential significance of the data. The initial choices of treating the main variables separately rather than jointly, and of using population subsetting rather than statistical controls, seems appropriate for a first exploration of the proposed model and of the measures employed, but not for a definitive testing of the data and model.

The results suggest that further inquiry may legitimately simplify the three main classes of variables--withdrawal, stress, and strain--by indexing or scaling. For example, the six strain measures, with the likely exception of the measure of work-related illness and injury, are all moderately intercorrelated and all appear to be functioning in analysis in a similar manner. This suggests that an index combining the set would be valid, would not conceal important differences among the component measures, and would provide the advantage of enhanced reliability and analytic simplicity. Similarly, the stress measures, with the possible exception of skill underutilization, might be combined into a common index of work role stress. The formation of a scale of withdrawal is more problematical, although the measures of absence, lateness, low involvement, and low interaction could be indexed, leaving job transfer and voluntary termination as separate variables.

With such a simplification of the variable matrix, and with the added confidence achieved as to the validity and reliability of such indices, analytic strategies would then be preferred which employ the whole of the available population rather than selected segments, and which treat all variables simultaneously along with some of the control variables (such as job tenure, occupationl status, and the like) which may have a bearing on the nature of an explanatory model.

References

Beehr, T. A. Role ambiguity as a role stress: Some moderating and intervening variables. Unpublished doctoral dissertation, The University of Michigan, 1974.

Beehr, T. A., Walsh, J. T., & Taber, T. D. Relationship of stress to individually and organizationally valued states: Higher order needs as a moderator. Journal of Applied Psychology, 1975, 61, 41-47.

Belloc, N., Breslow, L., & Hochstim, J. Measurement of physical health in a general population survey. American Journal of Epidemiology, 1971, 93, 328-329.

Caplan, R. D., Cobb, S., French, J. R. P., Jr., Harrison, R. V., & Pinneau, S. R., Jr. Job Demands and Worker Health. (NIOSH Series, Contract No. HSM-99-72-61) U. S. Government Printing Office, 1975.

French, J. R. P., Jr., & Caplan, R. D. Organizational stress and individual strain. In A. J. Marrow (Ed.), The Failure of Success. New York: AMACOM, 1973, 30-66.

Kasl, S. Work and Mental Health. New Haven: Yale University, 1973.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. Organizational Stress: Studies in Role Conflict and Ambiguity. New York: Wiley, 1964.

Kornhauser, A. Mental Health and the Industrial Worker. New York: Wiley, 1965

Margolis, B. L., Kroes, W. H., & Quinn, R. P. Job stress: An unlisted hazard. Journal of Occupational Medicine, 1974, 10, 659-661

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey. Ann Arbor: Institute for Social Research, 1974.

Chapter 17

TURNOVER: AN INQUIRY INTO THE EFFECTS OF JOB SATISFACTION, WORKING CONDITIONS AND QUALITY OF JOB PERFORMANCE

by

Terry A. Beehr

ABSTRACT

Three questions are addressed: (1) Does expressed intent to leave one's employer correspond to later behavior of actually leaving? (2) Does job (dis)satisfaction provide an early signal of an employee's likelihood of staying or leaving? (3) Does the quality of employment (working conditions) impact upon likelihood of leaving or staying? and (4) Do substandard working conditions and job dissatisfaction operate selectively to drive out the superior performers more than the less well performing employees?

The data employed in this analysis included measures of working conditions and job satisfaction at Time 1, along with measures of intention to leave. A follow-up measurement about 20 months later provided information on intention to leave at that later time as well as on instances of actual turnover, both voluntary and involuntary.

The results show that intention to leave, initial job dissatisfaction, and initial low quality of employment conditions all serve as leading indicators of actual voluntary turnover. It was found that these factors appear to operate in much the same way for both high performing employees and for those whose performance is judged less favorably.

Eighteen per cent of the initial population left their original place of employment within a period of 20 months. Ninety-two per cent of the leavers were classified as voluntary, and they included superior and inferior performers in about equal proportions. Some discussion is provided as to the costs and benefits associated with turnover, the persistence of turnover during periods of rising unemployment, and the implications for turnover reduction.

Chapter 17

TURNOVER: AN INQUIRY INTO THE EFFECTS OF JOB SATISFACTION, WORKING CONDITIONS AND QUALITY OF JOB PERFORMANCE

It is a common event for the American worker to leave his present job to seek one with another employer. The median term of employment with present employer, for a sample of employed adults in the United States, is about two years (Quinn, et al., 1971). National estimates of the total number of job changes per year vary according to the year in question and basis for estimation, but appear to be at the rate of about 50 million per year. The job changes, of course, are distributed in a far from random fashion, and under a variety of conditions that require different assessments. Many individuals, such as those in occupations that are seasonal, temporary, or designed for high turnover, may have frequent changes of employer. Others make a lifetime career with a single employer. Some switching is involuntary, as when the employing firm closes down; others are voluntary.

There has developed in recent years an interest in the cost and social efficiency of turnover. The costs and benefits accrue differently to the society, to the employer and to the individual and are unlike in different sectors of the economy. From a societal point of view, one can assume that there are general social benefits from inter-job mobility to the extent that it promotes career progress, promotes a competitive labor market, or promotes a ready inter-sector and inter-regional accommodation to economic changes.

From the employer's point of view there are inevitable out-of-pocket costs relating to the departure and replacement of employees. The lowest such turnover cost estimate known to us is $150 per instance of loss and replacement in a very efficient high-turnover organization. The typical cost may be as high as $2,000 if all costs are fully accounted. For the worker, the voluntary job switch can be presumed to be advantageous in most instances although the typical net cost or gain does not lend itself to meaningful calculation or estimate. The employer is no doubt advantaged if his poorly-performing employees leave.

There are accumulating some theories and some empirical data regarding turnover. The account that follows is intended to advance the matter. The specific concern is with the following questions: (1) Does expressed intent to leave one's employer correspond to the later behavior of leaving? (2) Does job satisfaction (dissatisfaction) provide an early signal of an individual's likelihood of staying or leaving? (3) Does quality of employment (working conditions) impact upon likelihood of leaving or staying? (4) Are the better-performing employees more responsive than the poorly-performing employees to job dissatisfaction and sub-standard working conditions when deciding to leave?

The context for this inquiry was a study of the life and job changes experienced by several hundred typical employed adults over a time span of about 20 months. One aspect of the study provided for measuring intentions to change employers, and a follow-up to ascertain whether there was in fact a job change. In addition, information was available regarding the quality of employment in the initial job, satisfaction with the initial job and its various facets, and the employer's opinions about the quality of performance of the individual in the initial job.

The growing literature relevant to such questions supports the merit of further study. Dissatisfaction with working conditions is surely a factor in voluntary job changes as shown in reviews by Herzberg, et al. (1957), Schuh (1967), Mangione (1972), and others. In the special case of socially disadvantaged new entrants into industrial employment, there appear to be strong effects of substandard working conditions upon voluntary turnover (Quinn, Fine, & Levitin, 1970). There is ample evidence that voluntary occupational mobility is associated with economic and social advantage to the movers (Lansing & Mueller, 1967). Some kinds of involuntary switching, at least, are advantageous to the workers, even though the prevalent experience may be otherwise (Kim, Roderick, & Shea, 1973).

From the employers' point of view, the crux of the matter is whether it is the better employees who are most likely to leave, and whether it is plausible on examination that the departure of better employees might be reduced through the provision of improved quality of employment and of more satisfying work environments.

Methods and Measures

The study population was not, strictly speaking, a sample of any defined segment of the workforce but a convenient and diverse population in typical non-managerial occupations. Information was obtained initially from and about 651 employees of five organizations in Michigan. Included were 173 members of a printing firm, 21 members of an R&D firm engaged in research in the physical and mathematical sciences, 120 and 124 members respectively in two automotive supply manufacturing firms, and 213 members of the service departments of a hospital. Supervisors were included. Details on these populations are described in Chapters 1 and 2.

The data pertinent to the present analysis were collected on two occasions, separated by approximately 20 months -- Fall, 1972 and Fall, 1974. While the study was designed as a panel study, to include re-measurement of the original population, there did occur some expected population loss, i.e., people who were no longer employed at the same place at Time 2. These population changes have relevance to the analysis described here because of this loss from the original Time 1 population. During the period between Time 1 and Time 2 measurements, the economy took a downturn with the effect that job change opportunities became less readily available and employed people generally became more cautious about voluntary termination from secure employment. Even so, during the period of about 20 months, 18 percent of the original Time 1 population changed employers, and of these, 92 percent described their change as voluntary or were reported by their original employer to have terminated voluntarily.

The study from which this analysis is derived involved a broad array of information as to domains of content and methods of measurement. These measures are described, together with technical specifications as to their psycnometric properties, in Chapters 1-3. The specific measures employed in the present analysis are described next. A list of these measures appears in Table 1 together with the number of cases, range, mean, standard deviation and coefficient alpha (an internal consistency index) for each measure.

Table 1

Means, Standard Deviations, and Ranges of All Variables

	Number of Cases	Range	Mean	Standard Deviation	Coefficient Alpha
Supervisor Ratings					
Time 1	286	1.2-7.0	5.48	1.00	0.89
Turnover					
Time 1 Turnover Intent	651	1.0-5.0	4.04	1.46	*
Time 2 Turnover Intent	267	1.0-5.0	1.82	1.40	*
Overall Turnover	310	1.0-2.0	1.24	0.43	*
Voluntary Turnover	286	1.0-2.0	1.17	0.38	*
Time 1 Satisfaction					
Facet-free	649	1.0-5.0	3.63	1.07	0.66
Challenge	647	1.0-4.0	3.05	0.68	0.84
Comfort	647	1.0-4.0	3.05	0.54	0.68
Resources	647	1.0-4.0	3.07	0.64	0.69
Pay	647	1.0-4.0	3.03	0.76	0.90
Time 1 Working Conditions					
Total	650	2.1-4.9	3.70	0.50	*
Challenge	650	1.3-5.0	3.68	0.84	*
Resources	637	1.2-5.0	4.05	0.76	*
Financial Rewards	648	1.5-5.0	3.54	0.61	*
Comfort	650	2.0-5.0	3.71	0.53	*

*Not applicable

Employee Performance Time 1 is represented by supervisory ratings as to the subordinate's quantity and quality of work, dependability, lateness, creativity at work, liking for the work, responsibility, and ability to get along with others. Ratings were made on 7-point scales in the polar semantic differential format, with the ratings averaged to obtain an overall index. For various practical reasons, these supervisory ratings were obtained on only 286 of the original Time 1 population of 651; some employers declined to provide the ratings, and some employees did not have a supervisor willing and qualified to provide such ratings, as in cases where the supervisor, or employee, was too new in his job.

Job Satisfaction Time 1 is represented by five indices, one being an overall index derived from a series of general interview questions evaluative of the job but not referring to any specific aspect or facet of the job. The remaining four indices of job satisfaction referred specifically to the following selected factors: Challenge (use of skills, opportunity to learn, etc.); Comfort and Convenience; Adequacy of Resources (materials, equipment, information); and Financial Rewards (pay, benefits). The derivation and technical properties of these indices of job satisfaction are described by Quinn & Shepard (1974).

Working Conditions Time 1 (quality of employment) was represented by five indices, all derived from responses in interview to a series of questions asking for "objective", non-evaluative descriptions of specific features of the respondent's job and work environment. The component facets match those used for the satisfaction measures (see above), but were introduced at a separated time in the course of a long interview. Indices were available for the four selected factors -- Challenge, Resource Adequacy, Financial Rewards, and Comfort -- and for an overall index comprising all of the factors. The derivation of these quality of employment indices is likewise described by Quinn & Shepard (1974) and in Chapter 3.

Turnover Intent, Time 1 and Time 2, is represented by the respondent's answer in interview to the following question: "Taking everything into consideration, how likely is it that you will make a genuine effort to find a new job with another employer within the next year -- very likely, somewhat likely, or not at all likely?" The same question was asked in the Time 2 interviews to ascertain intention to turnover at that later time.

Actual Turnover Time 2 was determined by a canvass of the employers' records approximately 20 months after the Time 1 measurements had been obtained. The instances of departure were all treated as turnover whether or not the individual actually had joined another employer. Some follow-up search was made to contact the departed employees, and to confirm the employers' classification of the event as "voluntary" or "involuntary". There were few discrepancies. Ninety-two percent of the instances were classified as voluntary. The remaining instances included discharge, retirement, reduction in force, etc.

The analysis strategies for these data were derived directly from the question stated in the introduction, and relied primarily upon correlational analysis. The questions are whether those employees who intend to leave actually do leave or continue to intend to leave, whether leaving is associated with job dissatisfaction or with poor quality of working conditions and which association is the stronger, and, finally, whether the high performance employees are more responsive than low performance employees to dissatisfaction and quality of working conditions when deciding to leave or to stay. Four population sub-sets were considered: the population at Time 1 were dichotomized into "high" and "low" categories according to the supervisors' ratings of their performance; and the employees leaving their employers (Time 2) were treated in some comparisons as "overall turnover" (all instances) while in others the "voluntary" instances are treated separately. The correlational tests involve the relationships between the Time 1 variables (intention, satisfaction, working conditions) and the Time 2 variables (still intend to leave, did leave, left voluntarily). Where feasible, significance tests were employed to assess the observed correlational or scale differences.

Results

Table 2 shows in the first numerical column the correlations between intention to leave (Time 1) and the turnover measures at Time 2. The correlations are all positive and strongly significant, and in a plausible pattern. Time 1 intention to leave correlates most strongly with actual voluntary departure, next most strongly with continuing intention to leave, and less strongly with "overall turnover" which includes some individuals leaving under duress, not by choice.

The relationships between initial job satisfaction, and initial (Time 1) quality of working conditions, on the one hand, and the measures of turnover on the other, are shown in Table 3. With respect to Time 1 turnover intention, the five concurrent indicators of job satisfaction are all significant correlates. The same Time 1 satisfaction measures, except for one, also correlate with measures of intent to leave at Time 2, about 20 months later. The correlations between job satisfaction at Time 1, and the actual later act of leaving are weaker, as might be expected, but still significant in

Table 2

Correlations of Time 1 Turnover Intent with Time 2 Turnover Variables

	Time 1 Turnover Intent			
	Total Sample	High Supervisor Ratings	Low Supervisor Ratings	Significance[a] of Difference between Correlations for High & Low Supervisor Ratings
Time 2 Turnover Intent	.35** (n=155)	.54** (n= 31)	.42** (n= 40)	p = .54
Overall Turnover	.28** (n=300)	.16 (n= 60)	.27** (n= 88)	p = .51
Voluntary Turnover	.39** (n=277)	.31* (n= 54)	.45** (n= 81)	p = .36

*p < .05

**p < .01

[a] Two-tailed tests

Table 3

Correlations of Time 1 Satisfactions and Working Conditions with Turnover Variables

	Turnover Intent		Actual Turnover	
	Time 1 (n=631)	Time 2 (n=155)	Overall (n=300)	Voluntary (n=279)
Satisfactions				
Facet-free	-.42**	-.26**	-.17**	-.20**
Challenge	-.34**	-.13	-.18**	-.22**
Resources	-.26**	-.21**	-.08	-.07
Pay	-.36**	-.24**	-.02	-.09
Comfort	-.27**	-.24**	-.14*	-.16**
Working Conditions				
Total	-.25**	-.11	-.07	-.10
Challenge	-.29**	-.06	-.12*	-.16**
Resources	-.17**	-.17*	-.01	-.04
Financial Rewards	-.09*	.05	.00	.00
Comfort	-.16**	-.18*	-.09	-.14*

*p < .05

**p < .01

over half of the tests for voluntary departures.

The comparable correlations between quality of working conditions at Time 1, and the four indicators of turnover are similarly patterned but weaker, with three (of ten) tests against actual turnover being significant.

The questions relating to the differential reactivity of the "high" and "low" performance employees to job satisfaction and to the quality of their working conditions is somewhat more complex than the other questions raised. Three tables are needed to display the evidence.

Table 2 shows the correlations between initial (Time 1) intention to leave and the subsequent decision to leave or continued intentions to leave, separately for the high and low performance employees. While the differences between "high and "low" performance employees in their responsiveness to their own Time 1 turnover intention are not statistically significant, there is a suggestive pattern that appears. The suggestion is that the low performance employees may be more likely to act on their initial intention to leave, while the high performance employees initially intending to leave may be more likely to stay on but to retain their intention to leave.

Table 4 allows the comparison of the "high performance" and "low performance" employees as to their standing on the turnover, satisfaction and working conditions variables. While there are consistent differences showing that the high performance employees reported more job satisfaction and better working conditions than the low performance employees, none of the differences is large and none significant. The "highs" also show small and non-significant differences as to intention to leave and actual turnover.

Table 4

Means and Standard Deviations of All Variables:
High versus Low Supervisor Ratings

	Supervisor Ratings	Number of Persons	Mean**	Standard Deviation
Turnover				
Time 1 - Intent	High	138	3.97	1.55
	Low	148	3.99	1.52
Time 2 - Intent	High	56	1.64	1.43
	Low	81	1.86	1.41
Overall	High	62	1.29	0.46
	Low	88	1.22	0.41
Voluntary	High	55	1.20	0.40
	Low	81	1.15	0.36
Time 1 - Satisfaction				
Facet-free	High	137	3.51	1.09
	Low	148	3.78	1.00
Challenge	High	138	3.02	0.69
	Low	148	3.8	0.61
Comfort	High	138	3.04	0.54
	Low	148	3.14	0.48
Resources	High	138	3.02	0.66
	Low	148	3.20	0.54*
Pay	High	138	2.98	0.77
	Low	148	3.09	0.76
Time 1 - Working Conditions				
Total	High	137	3.74	0.52
	Low	148	3.76	0.45*
Challenge	High	137	3.68	0.86
	Low	148	3.84	0.79
Resources	High	136	4.01	0.80
	Low	148	4.13	0.68*
Financial Rewards	High	137	3.57	0.60
	Low	148	3.64	0.54
Comfort	High	137	3.76	0.52
	Low	148	3.76	0.45*

*p (difference) <0.05

**Except for the turnover measures, low scores represent more of the named variable.

Table 5 displays the same information as appears in Table 3 but separately for the high and low performance employees. Time 1 job satisfaction measures and Time 1 working conditions measures are correlated with the four turnover indicators: intention to leave (Time 1 and Time 2), overall actual turnover and voluntary turnover. The table is to be read with the hypothesis in mind that employees rated by their supervisors as superior in performance would be found to be more reactive than the comparison group in the sense that low satisfaction and poor working conditions would more strongly be associated with an intention to leave and with actual departure.

With respect to intention to leave, the differences were uniformly in the predicted direction (two minor reversals) and nine of the 20 differences in size of correlation are significant. The differences were relatively small with respect to Time 1 versus Time 2 intention to leave because of opposite changes in the two compared groups. Among the high performance employees, Time 1 satisfactions and working conditions correlated with intention to leave less strongly at Time 1 than 20 months later, whereas among the low performance employees, the correlations at Time 1 all became insignificantly different from zero at Time 2.

With respect to actual turnover, voluntary and overall, the correlations tended to be somewhat stronger for the low performance employees than for the high performers (contrary to prediction), but none of the differences approached significance.

Discussion

These data replicate the finding of previous studies showing that an expressed intention to leave one's employer is related to the later behavior

Table 5

Correlations of Time 1 Satisfactions and Working Conditions with Turnover Variables: Employees with High versus Low Supervisor Ratings

	Time 1 Turnover Intent			Time 2 Turnover Intent			Overall Turnover			Voluntary Turnover		
	(r_1) High Supv. Rtgs. (n=134	(r_2) Low Supv. Rtgs. (n=148)	Significance^a of Difference between r_1 & r_2	(r_3) High Supv. Rtgs. (n=41)	(r_4) Low Supv. Rtgs. (n=40)	Significance^a of Difference between r_3 & r_4	(r_5) High Supv. Rtgs. (n=60)	(r_6) Low Supv. Rtgs. (n=88)	Significance^a of Difference between r_5 & r_6	(r_7) High Supv. Rtgs. (n=54)	(r_8) Low Supv. Rtgs. (n=81	Significance^a of Difference between r_7 & r_8
Satisfactions												
Facet-free	-.46**	-.35**	p=.14	-.61**	-.19	p=.02	-.11	-.17	p=.64	-.23	-.25*	p=.55
Challenge	-.25**	-.24**	p=.46	-.35	.05	p=.05	-.13	-.26*	p=.79	-.23	-.31**	p=.68
Resources	-.35**	-.15	p=.04	-.59**	-.01	p=.00	-.13	-.04	p=.30	-.25	-.04	p=.12
Pay	-.33**	-.40**	p=.62	-.48**	-.12	p=.05	.00	.06	p=.36	-.12	-.08	p=.29
Comfort	-.25**	-.12	p=.13	-.30	-.17	p=.29	-.09	-.10	p=.52	-.15	-.21	p=.63
Working Conditions												
Total	-.27**	-.09	p=.02	-.37*	-.15	p=.17	.03	-.06	p=.70	-.05	-.15	p=.71
Challenge	-.24**	-.21*	p=.40	-.30	.09	p=.02	.02	-.14	p=.83	-.07	-.27*	p=.87
Resources	-.29**	.01	p=.00	-.56**	.09	p=.00	.04	.02	p=.55	.01	.07	p=.37
Financial Rewards	-.11	-.04	p=.28	-.09	.17	p=.15	.15	.05	p=.72	.17	.07	p=.71
Comfort	-.18*	-.04	p=.12	-.28	-.14	p=.29	-.19	-.13	p=.36	-.23	-.28**	p=.62

*p < .05
** < .01
^{a}One-tailed tests

of actually leaving. The data add the confirming touch of showing that this association between intent and subsequent action is somewhat stronger when tested for those whose leaving is classified as voluntary rather than involuntary.

A role for job satisfaction in the initiation of turnover intention and action is made plausible by these data, as job (dis)satisfaction at Time 1 correlates significantly with intention to leave both concurrently and also at a later time, and job satisfactions correlate significantly as well (although less strongly and less significantly) with actual decisions to leave one's employment.

Inferior working conditions appear to figure more weakly than job satisfaction, and perhaps less immediately, in the induction of intention to leave and decision to leave. The correlational evidence shows working conditions at Time 1 to be related to concurrent intention to leave, less strongly with later intention to leave, and weakly although significantly with actual decisions to leave.

One can speculate that there may be operating a causal linkage system in which inferior working conditions induce low job satisfaction (an established relationship of considerable strength as shown by Quinn, et al., 1971), with dissatisfaction then entering into the behavioral decision to leave. However, such a linkage can not be verified with the present data.

The major question addressed in this inquiry is whether turnover tends to occur disproportionately among exactly those employees that employers would most like to keep -- i.e., those rated most favorably by their supervisors -- and whether such superior and potentially superior employees would be found more ready to convert experiences of poor working conditions and low job satisfaction into the decision to terminate. The evidence is mixed and complex. As to absolute rates of turnover, the highly-rated employees are slightly but non-significantly more likely to leave than the low rated employees. But, when quality of working conditions and job satisfaction are taken into account, the evidence runs against the hypothesis that high-performing employees are more reactive to their job environments. The correlations between satisfactions and quality of working conditions, used as predictors, and actual voluntary turnover are somewhat higher, on average, for the low-rated employees and more frequently significant, while none of the correlations for the highly-rated employees reached significance. Contrarily, as to intention to leave, the higher-rated employees appeared to be more responsive than the low-rated employees to poor working conditions and low job satisfaction. There is an anomaly: a poor employment situation appears to increase the intention of highly-rated employees to leave (relative to the low-rated employees) but the reverse is true for their actual behavior of leaving. In sum, for the population and time period in question, there is no evidence that relatively poor conditions of employment operate selectively to drive out the better employees more than the poorer ones.

Some minor cautions are to be noted in relation to this conclusion. First, there is reason to be wary of the classification of turnover events as "voluntary" or "involuntary", even when both parties concur in judgement, for it is likely that some employees "choose" to leave in response to various signals that their departure would be welcomed even though not required. Second, the supervisory ratings and the employee's intention-to-leave measures may be thought to be somewhat contaminated even though from seemingly independent sources, as in the case where the supervisor is aware of the subordinate's inclination to stay or leave and incorporates this knowledge in his ratings; however, this suspicion is allayed by the absence of significant mean differences in variable scores (Table 4) between high and low rated employees. Third, while there is probably some spurious increment, because of common times and measurement methods, to the correlations between intention to leave, job satisfaction and working conditions at Time 1, it is to be noted that the effect of this is not strong enough to suppress the emergence of systematic changes in relationships between Time 1 and Time 2.

One might think, further, that the results of this analysis may be influenced by exogenous factors inadvertently introduced. This could happen through the supervisory ratings being biassed with respect to the subordinates' characteristics in a way that obscures the relationships tested in the analysis. However, the demographic differences between the highly rated and low rated employees proved upon examination to be generally small and non-significant, and wholly uninterpretable in this context.

In summary, the evidence is (1) that employees who intend to leave are indeed likely to do so. (2) that exposure to inferior working conditions and job dissatisfaction do predispose employees to both intention to leave and actual departure, and (3) with respect to the inducement of turnover intention and actual leaving, the effects of prior poor working conditions and dissatisfaction are about the same for superior employees and those judged inferior.

References

Beehr, T. A. Perceived situational moderators of the relationship between subjective role ambiguity and role strain. *Journal of Applied Psychology*, 1976, *61*, 35-40.

Herzberg, R., Mausner, B., Peterson, R., & Capwell, D. Job attitudes: Review of research and opinion. Pittsburgh: Psychological Services of Pittsburgh, 1957.

Kim, S., Roderick, R., & Shea, J. *Dual Careers: A Longitudinal Study of the Labor Market Experience of Women*, Volume 2. Washington, D.C.: United States Government Printing Office, 1973.

Kraut, A. I. Predicting turnover of employees from measured job attitudes. *Organizational Behavior and Human Performance*, 1975, *13*, 233-243.

Lansing, J. B., & Mueller, E. *The Geographic Mobility of Labor*. Ann Arbor: Institute for Social Research, 1967.

Mangione, T. W. Turnover: Some psychological and demographic correlates. In R. P. Quinn & T. W. Mangione (Eds.), *The 1969-70 Survey of Working Conditions: Chronicles of an Unfinished Enterprise*. Ann Arbor: Survey Research Center, 1973.

Mangione, T. W. Turnover: A model and a review of the literature. Ann Arbor: Survey Research Center, 1972.

Quinn, R. P., Seashore, S. E., Kahn, R. L., Mangione, T. W., Campbell, D. B., Staines, G. L., & McCullough, M. R. Survey of Working Conditions: Final Report on Univariate and Bivariate Tables. U.S. Government Printing Office, 1971.

Quinn, R. P., Staines, G. L., & McCullough, M. R. Job Satisfaction: Is There a Trend? Washington, D.C.: U. S. Department of Labor, 1974. Manpower Research Monograph No. 30.

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey: Descriptive Statistics, with Comparison Data from the 1969-70 Survey of Working Conditions. Ann Arbor: Survey Research Center, 1974.

Quinn, R. P., Fine, B. D., & Levitin, T. E. Turnover and training: A social psychological study of disadvantaged workers. Springfield: National Technical Information, Operations Division, Document #: PB194775, 1975.

Schuh, A. J. The predictability of employee tenure: A review of the literature. Personnel Psychology, 1967, 20, 133-152.

Chapter 18

SUPERVISOR-SUBORDINATE SIMILARITY: A DETERMINANT OF SUBORDINATE RATINGS AND REWARDS

by

Veronica Feria Nieva

701

ABSTRACT

It has been asserted that supervisors tend to allocate more favorable performance ratings, and more generous rewards, to those subordinates who are most like themselves, and that this tendency holds also for characteristics that are not obviously pertinent to job qualifications or job performance. This proposition was examined within a varied sample of workers and their supervisors in five work organizations.

The similarities considered included five demographic variables, five job relevant attitudes, and four measures of personality. Significant relationships were found between supervisor-subordinate similarity and favorableness of ratings and rewards, and they occurred consistently across all types of similarity measures. A test for possible moderating effects arising from job certainty and the quality of superior-subordinate interactions showed that the relationships were stronger for workers in jobs of relatively low certainty; the moderating effects of superior-subordinate interaction quality were very weak. Availability of follow-up measures after a period of about twenty months allowed investigation of the probable direction of causation; the data suggested that: (1) similarities tend to cause favorable ratings; (2) favorable ratings tend to cause favorable rewards; and (3) favorable rewards tend to induce further similarities, perhaps by mutual selection and/or modification of attributes.

In the case of ratings, the effects of superior-subordinate similarity were somewhat stronger for female subordinates and for female superiors, compared with males, and stronger for black subordinates and superiors, compared with whites. In the case of rewards, the reverse was found, with stronger relationships for males and whites compared with their counterparts.

The results are interpreted as a further explication of the organizational circumstances and mechanisms through which non-rational factors may affect the allocation of rewards and approbation in work organizations.

Chapter 18

SUPERVISOR-SUBORDINATE SIMILARITY:
A DETERMINANT OF SUBORDINATE RATINGS AND REWARDS*

In an ideal bureaucratic world, only the chracteristics of an employee, assessed according to the organization's well-specified and universally applied standards, would be relevant to the evaluations made of, and the subsequent rewards allocated to, that individual. For the most part, these standards would conform to the achievement ideology, which legitimates the use of criteria of merit or effective performance, however defined (Levitin, Quinn, and Staines, 1971). The emphasis on achievement-related norms, in combination with bureaucratic ideals of impersonality and objectivity, creates expectations that organizational processes such as personnel evaluation and reward allocation are carried out rationally in accordance with stated policies and procedures.

Unfortunately this ideal state is achieved none too often. Between the statement of the organization's rational official policy and its concrete implementation often lies a chasm. Research in recent years has shown that factors other than ability or achievement-related characteristics of job holders significantly influence personnel decision-making in organizations. Representative non-ability factors found to affect organizational judgmental processes include personal appearance (Carlson, 1967; Dipboye, Fromkin and Wiback, 1975), sex-role stereotypes (Deaux and Emswiller, 1974; Cecil, Paul and Olins, 1973; Rosen and Jerdee, 1973, 1974), race, sex, and national origin (Bowman, 1964; Hamner, Kim, Baird and Bigoness, 1974). Assumptions about the objective rationality of such processes are, therefore, frequently unwarranted.

---

*This brief report is abstracted from the comprehensive source document (Nieva, 1976). This analysis of the project data was made possible by Dissertation Grant No. 91-26-75-37 from the Employment and Training Administration, Department of Labor.

This study can be categorized among a number of others that have tried to investigate some of the unwritten laws that govern personnel decisions in reality in comparison to the rules that are formally stated by organizations. The theoretical backing of the research is taken from the social-psychological study of interpersonal similarities and attraction which has demonstrated conclusively that similarities on a wide range of dimensions influence favorable affect across a wide range of content in both laboratory and field situations. Within the work setting, recent studies have demonstrated the effect of similarities on internal organizational processes, e.g., voting choice, loaning money and selection recommendations. The general proposition of the study is based on this similarity-attraction relationship. The study assesses the relationships between various types of supervisor-subordinate similarity and the ratings and rewards received by the subordinate.

Data Collection and Instruments

The data used in this study were collected in two phases. Five organizations -- a hospital, a printing company, two automotive supply companies, and a research and development firm -- participated in the first phase. Phase II data were collected from the hospital and the two automotive supply companies after an interval of about two years.

The sample base for this study included 649 individuals, but the several analytic steps involved lesser numbers, as not all individuals were provided with supervisor ratings, and not all were included in both data collection phases.

Three types of instruments were used to gather data for the study. Interviews covering a variety of content areas, e.g., job attitudes, physical and mental health, and various personality indices, were conducted in the respondents' homes. In addition, all supervisors were asked to rate each subordinate on a 7-point semantic differential rating scale along eight dimensions. Finally, personnel records were collected from the organization.

Further information about the sample and measures is available in the source document (Nieva, 1976) as well as the present report chapters on data acquisition methods. Table 1 shows in further specific detail the variables employed and their respective sources.

Conceptual Framework and Hypotheses

The conceptual framework that guided the design of this investigation is displayed in Figure 1. The figure distinguishes among panels of

Table 1

Variables Used in this Study

Variables	Number of Items	Source
Independent Variables		
Similarity in:		
A. Demographic Data		
1. Age	1	Interview
2. Education	1	Interview
3. Race	1	Interview
4. Sex	1	Interview
5. Marital status	1	Interview
B. Job-related Attitudes		
1. Value of intrinsic job factors	7	Interview
2. Value of extrinsic job factors	2	Interview
3. Value of interpersonal job factors	4-3*	Interview
4. Overall job satisfaction	4	Interview
5. Job involvement	4-3*	Interview
C. Personality		
1. Tolerance of Ambiguity	3	Interview (card sort)
2. Self-Esteem	3	Interview (self-administered questionnaire)
3. Locus of Control	3	Interview (self-administered questionnaire)
4. Need for Affiliation	3	Interview (self-administered questionnaire)

continued...

---

*Number of items in Phase II scale

**Collected from records in Phase I and from the interview in Phase II.

Table 1 continued

Variables Used in this Study

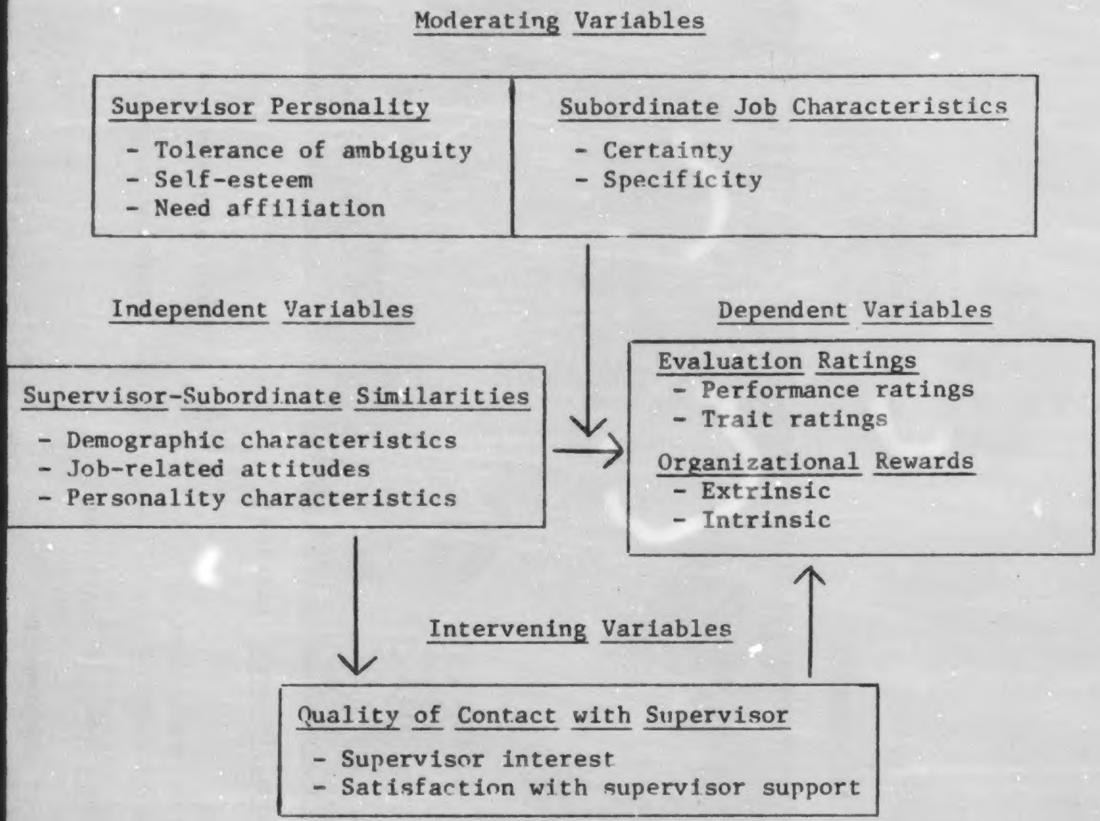
Variables	Number of Items	Source
Moderating Variables		
A. Job Characteristics		
1. Job Certainty	4	Interview (card sort)
2. Job Specificity	2	Observation
B. Supervisor Personality		
1. Tolerance of Ambiguity	3	Interview (card sort)
2. Need for Affiliation	3	Interview (self-administered questionnaire)
3. Self-Esteem	3	Interview (card sort)
Dependent Variables		
A. Evaluation Ratings		
1. Performance rating	2	Supervisor rating form
2. Trait rating	3	Supervisor rating form
B. Organizational Rewards (extrinsic)		
1. Pay level	1	Records/Interview**
2. Quality of employment: Financial Rewards	5	Interview
C. Organizational Rewards (intrinsic)		
1. Quality of employment: Challenge	12	Interview

---

*Number of items in Phase II scale.
**Collected from records in Phase I and from the Interview in Phase II.

Figure 1

Conceptual Framework of the Relationships between Supervisor-Subordinate Similarity and Subordinate Ratings and Rewards


Moderating Variables
Supervisor Personality
- Tolerance of ambiguity
- Self-esteem
- Need affiliation
Subordinate Job Characteristics
- Certainty
- Specificity
Independent Variables
Supervisor-Subordinate Similarities
- Demographic characteristics
- Job-related attitudes
- Personality characteristics
Dependent Variables
Evaluation Ratings
- Performance ratings
- Trait ratings
Organizational Rewards
- Extrinsic
- Intrinsic
Intervening Variables
Quality of Contact with Supervisor
- Supervisor interest
- Satisfaction with supervisor support

independent, dependent, moderating, and intervening variables. In general, the model proposes that the degree of similarity between supervisor and subordinate on a number of characteristics is positively related to the evaluation ratings and organizational rewards received by the subordinate. Three types of similarity (viz., demographic, attitudinal, and personality), two ratings and three reward measures are studied.

Similarity, as used in the study, is "actual", as opposed to "assumed" or "perceived". Indices of similarity are calculated from employee responses; no questions directly asking about perceptions of similarity were asked. The magnitude of the discrepancy between supervisor and subordinate on any given dimension is seen as primarily responsible for the relationship to the ratings and rewards received by the subordinate; thus, subordinate outcomes are expected to be most favorable when the supervisor-subordinate discrepancy is zero, and these outcomes are expected to be less favorable as the magnitude of the differences between supervisor and subordinate (either direction) increases.

In addition, the model proposes that the similarity-favorability relationships are moderated by personality characteristics of the supervisor. It is expected that the relationships between similarity and favorability would be stronger when the supervisor had lower self-esteem, less tolerance for ambiguity, and stronger needs for affiliation. The model also proposes that similarity-favorability relationships are moderated by characteristics of the subordinate's job. It is expected that jobs which are more clearly defined would leave less room for the influence of subjective factors, and thus minimize the relationships between similarity and favorability.

Finally, the model proposes that the quality of subordinate contact with the supervisor intervenes in the similarity-favorability relationship; i.e., part of the effect of supervisor-subordinate similarity (on the ratings and rewards received by the subordinate) operates indirectly through its effect on the quality of the relationship between supervisor and subordinate.

In addition to the relationships expressed in the model the study investigates two other problem areas. First, it explores the effects of race and sex on the general similarity-favorability relationship. In this regard, the following questions were addressed:

1. Does the relationship between similarity and favorability exhibit the same patterns for black and white, male and female subordinates?
2. Are there differences in susceptibility to the similarity-favorability relationship between black and white, male and female supervisors?
3. Do these relationships exhibit the same patterns when the supervisor and subordinates are of the same/different race, or when they are of the same/different sex?

Second, it tries to determine whether the proposed similarity-favorability relationships are causal and the direction of the causality; i.e., does similarity cause subordinate ratings and rewards more strongly than ratings and rewards cause similarity? Within the laboratory setting, the causal influence of similarity on favorability of response has been repeatedly documented; however, causality has been difficult to support in most field situations which have been largely confined to correlational data collected at one point in time. The panel data used in this study make it possible to investigate the direction and magnitude of causal relationships in the field.

Summary of Results

Direct Relationships

In support of the general hypothesis, the study found that the greater the absolute value of the similarity between supervisor and subordinate, the more favorable were the ratings and rewards received by the subordinate. The general hypothesis was approached on two levels: (1) The individual relationships between each similarity and each subordinate rating and reward (not shown here but available in the source document); and (2) The cumulative relationships between the similarities considered together and the subordinate outcomes (Table 2). With one exception, the relationships between each of the three classes of supervisor-subordinate similarities and each of the five subordinate outcomes are statistically significant. The relationships between all the similarities considered together and the subordinate outcomes are also significant. Similarities accounted for greater amounts of variance in the ratings and intrinsic rewards, the outcomes most directly influenced by the supervisor, and smaller amounts of variance were accounted for in extrinsic rewards, the subordinate outcome over which the supervisor had least control.*

Moderated Relationships

It was hypothesized that characteristics of the subordinate's job and of the subordinate's supervisor would moderate the relationships between the various similarities and subordinate outcomes. In support

---

* Additional analyses, not treated here but described in the source document, indicate that supervisor-subordinate similarities have effects on subordinate outcomes over and above the effects of the supervisor and subordinate traits.

Table 2

Multiple Correlations between Classes of Similarities and Subordinate Outcomes

		Performance Rating	Trait Rating	Pay	Financial Rewards	Intrinsic Rewards
Demographic similarities	Adj. R	.233**	.223**	.229**	.125*	.224**
	Adj. R^2	.054	.050	.052	.016	.050
Attitudinal similarities	Adj. R	.248**	.345**	.248**	.184**	.330**
	Adj. R^2	.061	.119	.061	.034	.109
Personality similarities	Adj. R	.176**	.123 n.s.	.140**	.175**	.198**
	Adj. R^2	.031	.015	.019	.031	.039
All similarities	Adj. R	.380**	.405**	.350**	.253**	.407**
	Adj. R^2	.144	.164	.122	.064	.166

*p <.05
**p <.01

of the hypothesis, stronger similarity-favorability relationships were found among holders of jobs that had lower job certainty than among holders of jobs that had greater job certainty (Table 3). However, no support was found for the hypothesis that job specificity would moderate the similarity-favorability relationship. Neither was support found for the set of hypotheses regarding the moderating effects of supervisor personality characteristics on the basic relationships. The supervisor's self-esteem, need for affiliation, and tolerance of ambiguity did not effectively moderate any of the relationships.

Intervening Relationships

It was also hypothesized that the quality of supervisor-subordinate contact would, in part, intervene between similarities and subordinate outcomes, i.e., part of the effect of supervisor-subordinate similarity (on favorability of subordinate ratings and rewards) operates through the quality of the subordinate's contact with the supervisor. On the whole, the data tended to support this proposition. Similarities were related to the proposed intervening variables, the latter were significantly related to the subordinate outcomes, and the partial correlations obtained when the proposed intervening variables were held constant were generally smaller than were the direct relationships between similarities and subordinate outcomes. However, because the differences between the zero-order and the partial correlations were slight, it was concluded that the intervening effects found were minimal.

Table 3

Multiple Correlations between Each Class of Similarities and Subordinate Outcomes Moderated by Job Certainty

		Performance Rating		Trait Rating		Pay		Extrinsic Rewards		Intrinsic Rewards	
		Job Certainty		Job Certainty		Job Certainty		Job Certainty		Job Certainty	
		Low	High	Low	High	Low	High	Low	High	Low	High
Demographic similarities	Adj. R	.213*	.187	.339**	.048	.333**	.165	.189*	.075	.386**	.133
	Adj. R^2	.045	.035	.115	.002	.111	.027	.036	.005	.149	.018
Attitudinal similarities	Adj. R	.261*	.203	.308**	.356**	.227*	.263*	.177*	.187*	.386**	.286**
	Adj R^2	.068	.041	.095	.127	.052	.069	.031	.035	.149	.082
Personality similarities	Adj. R	.238*	.000	.228**	.000	.206*	.081	.244**	.181*	.251**	.179*
	Adj. R^2	.056	.000	.052	.000	.042	.006	.060	.034	.063	.032
All similarities	Adj. R	.396**	.260	.464**	.313*	.422**	.305**	.308**	.254	.483**	.354*
	Adj. R^2	.157	.067	.215	.098	.178	.093	.094	.064	.234	.125

* $p < .05$
** $p < .01$

Race and Sex Analyses

The results of the analyses of race and sex effects show different results for ratings and for rewards. In the case of ratings, stronger relationships between similarity and favorability were found among female subordinates than among male subordinates; similarly, stronger relationships were found among subordinates with female supervisors than among subordinates with male supervisors. Parallel results were found for the race comparisons. Stronger relationships were found among black subordinates than among white subordinates; also, stronger relationships were found among those with black supervisors than among those with white supervisors. The results also showed that the similarity-favorability relationships as to rewards were stronger when supervisors and subordinates were of different sex or race than when they were of like sex or race.

Causal Analyses

Cross-lagged panel analyses were conducted to address the general question, "Does similarity cause ratings and rewards more than ratings and rewards cause similarities between supervisors and subordinates?" These analyses were conducted for combinations of attitudinal and personality similarity and each subordinate outcome. In line with expectations, it was found that similarities tended to cause favorable ratings. However, contrary to expectations, the overall pattern of results tended to show that similarity was caused by rewards rather than vice-versa. A connecting link was therefore proposed and tested; support was found for the suggestion that rewards would be shown as being caused by ratings. A chain effect was therefore proposed, leading from similarities to ratings, from ratings to rewards, and from rewards to further similarities (Figure 2.). The methods used and supporting data may be found in the source document.

Implications

Homogeneity: Its Possible Effects. The similarity-favorability relationships found in this study can be viewed in the light of forces towards increased homogeneity within organizations and in the larger society. To a certain extent, these pressures are part of the operation of organizational and societal socialization processes, necessary for the continued functioning of all social systems. Schein (1968) defines socialization as "the process of being indoctrinated and trained, the process of being taught what is important in an organization or subunit thereof". In this particular study, it appears that the subordinate learns that part of what is important is to become like one's superiors in the organization; subordinates become increasingly similar to their supervisors in response to the favorable ratings and rewards received from the supervisor. From the subordinate's point of view, the process of increasing similarity in response to favorable conditions can be interpreted in terms of expectancy theory (Vroom, 1964; Porter and Lawler, 1968), which postulates that an individual will behave in a certain fashion to the extent that such behavior is seen as leading to desirable outcomes. From another stream of research, Simons, Berkowitz and Moyer (1970) similarly propose that change towards the position held by another depends on the extent to which interpersonal similarity is perceived as instrumental

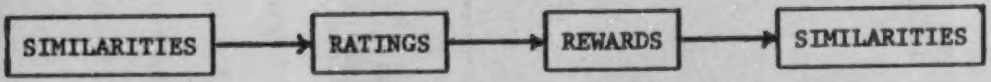

SIMILARITIES
RATINGS
REWARDS
SIMILARITIES

Figure 2

Causal Linkages Connecting Similarities, Ratings and Rewards in achieving that individual's ends.

The effects of these pressures for homogeneity may occur at the level of the individual, the organization, or the society as a whole. In more practical terms, the question is raised regarding the effects of homogeneity on the productivity of individuals, contrasting the possible positive effects of greater effectiveness due to more efficient communication (Triandis, 1959) with homogeneity's possible detrimental effects, particularly on creative and innovative work (Pelz and Andrews, 1966).

In a more humanistic vein, the effects of pressure towards homogeneity on the individual as a person are of interest. Because of the demands for stability and predictability that exist in most bureaucratically oriented organizations, organizations tend to focus on increasing the similarity and standardization among organizational members, to the extent that some theorists have viewed organizational interests as typically stunting to the individual employee's development. The achievement of a comfortable balance between rigid uniformity and complete anarchy, a desirable middle ground of "creative individualism" (Schein, 1968) appears to be generally problematic. Such an ideal condition would merge individual and organizational interests and would allow the flexibility for innovation and change that is needed by the organization in the long run.

There are also implications of the similarity effect for ethical and societal issues. For example, there is reinforcement of the notion that controllers of resources tend to dole them out to similar others. An interesting finding of the study is the magnification of the effects of similarity on the ratings provided by members of non-standard and relatively powerless groups, i.e., females and blacks, suggesting that greater conformity pressures are exerted on members of these groups. Extrapolating from these results obtained within organizational boundaries to the society at large, the similarity-favorability phenomenon can be interpreted as one of society's mechanisms of control. The tendency of administering limited goods to others who are like one's self can be an effective way of perpetuating the status quo. It tends to support the present demographic distribution of rewards, allowing exceptions for members of the non-standard groups who possess characteristics most closely resembling members of the majority group.

Strategies for Minimizing the Effects of the Similarity-Favorability Phenomenon. Recent studies have suggested strategies to increase accuracy of ratings similar to those used in this study. Rating theorists consistently warn against the fallibility of global types of ratings and have designed various devices, e.g., the critical incident technique and behaviorally anchored rating scales, to minimize subjective effects such as those that have been found in this study. Borman (1975) demonstrated that a training session introducing raters to halo error was effective in reducing the halo effect significantly in performance ratings. Similarly, other studies (Wexley, Sanders, & Yukl, 1973; Wexley & Nemeroff, 1974; Latham, Wexley, & Pursell, 1975) have found that intensive training workshops using a variety of methods have led to significant behavior change and reduced rating errors among experienced raters in organizations. The latter two studies are especially relevant in that they involved training focused on the similarity effect. They found that similarity accounted for almost no variance at all in the ratings made by the trained raters, in contrast to results for a control group. These studies give hopeful indications that the effects of different sources of error can, in fact, be minimized by such training efforts.

However, the interpretation must be made with caution that the relationship between similarities and favorability of outcomes for subordinates reveals only a lack of supervisor objectivity. An alternative explanation could be proposed: that similarities do, in fact, cause more effective performance, which is reflected in the higher ratings and rewards received by the subordinates. Homogeneity in groups has been found to relate to effective communication (Triandis, 1959); presumably, sharing basic assumptions and premises eliminates many steps and facilitates understanding in the communication process. It is highly probable that similarities between supervisor and subordinate carry both the illegitimate as well as some such legitimate effects.

/mm
Rev. 12.8.77

References

Borman, W. Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 1975, 60, 555-560.

Bowman, G. W. What helps or harms promotability? Harvard Business Review, 1964, 6-27.

Carlson, R. E. Selection interview decisions: The relative influence of appearance and factual written information on an interviewer's final rating. Journal of Applied Psychology, 1967, 51, 461.

Cecil, E. A., Paul, R. J., & Olins, R. A. Perceived importance of selected variables used to evaluate male and female job applicants. Personnel Psychology, 1973, 26, 397-404.

Deaux, K., & Emswiller, T. Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. Journal of Personality and Social Psychology, 1974, 29, 80-85.

Dipboye, R. L., Fromkin, H. L., & Wiback, K. Relative importance of applicant sex, attractiveness, and scholastic standing in evaluation of job applicant resumes. Journal of Applied Psychology, 1975, 60, 39-43.

Hamner, W. E., Kim, J. S., Baird, L., & Bigoness, W. J. Race and sex as determinants of ratings by potential employers in a simulated work sampling task. Journal of Applied Psychology, 1974, 59, 705-711.

Latham, G. P., Wexley, K. N., & Pursell, E. D. Training managers to minimize rating error in the observation of behavior. Journal of Applied Psychology, 1975, 60, 550-555.

Levitin, T. E., Quinn, R. P., & Staines, G. L. Sex discrimination and the American working woman. American Behavioral Scientist, 1971, 15(2).

Nieva, V. F. Supervisor-subordinate similarity: A determinant of subordinate ratings and rewards (Doctoral dissertation, The University of Michigan, 1976). Dissertation Abstracts International, 37, 6, B. (University Microfilms No. 76-27558)

Pelz, D. C., & Andrews, F. M. Scientists in organizations. New York: Wiley & Sons, Inc., 1966.

Porter, L. W., & Lawler, E. E. III. Managerial attitudes and performance. Homewood: Irwin, 1968.

Rosen, B., & Jerdee, T. H. Influence of sex role stereotypes on evaluations of male and female supervisory behavior. Journal of Applied Psychology, 1973, 57, 44-48.

Rosen, B., & Jerdee, T. H. Influence of sex role stereotypes on personnel decisions. Journal of Applied Psychology, 1974, 59, 9-14.

Schein, E. H. Organizational socialization and the profession of management. Industrial Management Review, 1968, 9, 1-16.

Simons, A. W., Berkowitz, N. N., & Moyer, R. V. Similarity, credibility and attitude change: A review and a theory. Psychological Bulletin, 1970, 73, 1-16.

Triandis, H. C. Cognitive similarity and interpersonal communication in industry. Journal of Applied Psychology, 1959, 43, 321-326.

Vroom, V. H. Work and motivation. New York: Wiley, 1964.

Wexley, K. N., & Nemeroff, W. Effects of racial prejudices, race of applicant, and biographical similarity on interviewer evaluations of job applicants. Journal of Social and Behavioral Sciences, 1974, 20, 66-78.

Wexley, K. N., Sanders, R. E., & Yukl, G. A. Training interviewers to eliminate contrast effects in employment interviews. Journal of Applied Psychology, 1973, 57, 233-236.

Chapter 19

ALTERNATIVE MOTIVATIONAL STRATEGIES: AN ASSESSMENT OF COSTS AND BENEFITS

by

John R. Klesh and Cortlandt Cammann

ABSTRACT

This paper assesses the compatibility between gains in the quality of work life experienced by workers and managerial interests in assuring high levels of work effort and involvement on the part of employees. Three strategic approaches to employee motivation are distinguished: contingent extrinsic rewards, contingent intrinsic rewards, and rules control. The impact of these different organizational strategies for motivation and control on employees' effort, involvement, work-related depression, and satisfaction was assessed using canonical analyses. In addition, the moderating effects of employee's need, race, sex, job autonomy, resource adequacy, job interdependence, and skillfit were explored.

The results indicated that, in general, the contingent extrinsic reward strategy and contingent intrinsic reward strategy are used together and these strategies are not used in conjunction with the rules control strategy. The first two approaches were associated with positive employee outcomes, while the rules control strategy was related to negative outcomes.

Finally, the major effects of the moderating variables were in the strength of relationships and not in their direction.

These results suggest that the contingent reward allocation strategies are mutually compatible and that neither involves a tradeoff between outcomes valued by employers and employees. The rules control strategy on the other hand is incompatible with the other two strategies both in terms of its use and its association with the outcome variables.

Chapter 19

ALTERNATIVE MOTIVATIONAL STRATEGIES: AN ASSESSMENT OF COSTS AND BENEFITS

This paper addresses the issue of the compatibility between gains in the quality of work life experienced by workers and managerial interests in assuring high levels of work effort, involvement, and commitment on the part of their employees. Three motivational strategies are distinguished which are selectively emphasized by different employers out of their beliefs concerning the most effective means of inducing effort and job involvement. These alternative strategies are compared with respect to their association with those outcomes that are likely to be favored by employers and those favored by employees. The central question is whether unavoidable tradeoffs exist such that an increment in the quality of working life is likely to imply a decrement in the employees' contribution of work effort and job involvement.

Beliefs concerning the incompatibility of work efficiency and quality of working life are endemic to modern society and permeate the recent history of employee-employer relationships. Many of the major gains in quality of working life such as those relating to job rights, minimum wages, unemployment compensation, and workplace safety, have been introduced by coercive means, legislation, or bargaining, rather than by voluntarism. This has been justified on the grounds that gains in quality of working life involve onerous costs to the employer which need to be balanced against desired gains for society and the employee. More recently, many proposed changes intended to improve the quality of working life have been such that they do not require a balancing of conflicting interests. On the contrary, they allow for differences of opinion as to whether a proposed change in the condition of working life may lead to a gain for the employee but a loss for the employer, or vice versa. Thus, at the present time, different labor unions and their leaders hold opposite views, as do employers, about the threats and potential gains inherent in various current approaches to the enhancement of working life.

One such domain of quandary concerns the strategy that an employer might emphasize in seeking to maximize effective employee work effort and commitment to the job. Two recent major reviews of the available evidence (Katzell, et al., 1975; Srivastva, et al., 1975) agree that the answer is far from clear or consistent, although there appear to be important convergences in some aspects of the problem. Both reports suggest that the effects of a given approach to work motivation appear to be contingent upon aspects of the work technology and job design, upon the personal attributes of the particular employee population, and upon factors that make up the context or environment in which the work is done. To illustrate the contingent nature of the work motivation strategies, there exist documented instances in which performance effectiveness was enhanced by detailed specification of expected work activity and quality controls, even though these are known commonly to have negative effects for the employees. Similarly, there are documented instances in which the redesign of jobs and work environments to optimize the quality of working life have failed to generate the expected gains in performance effectiveness.

The relevant theoretical positions are clear, and can be characterized as follows: (1) Most work is inescapably and inherently onerous, so that inducement of adequate performance depends upon providing generous pay, job security, and minimized ego involvement in exchange for performance of highly specified work activity (e.g., Fein, 1976); (2) There are both attractive and unattractive features of work and an equitable and satisfactory balance of performance and desired benefits will be achieved if extrinsic rewards are contingent upon performance (e.g., Lawler, 1973); (3) Work activities are, or can be made to be, intrinsically rewarding so that high performance can be elicited at low cost to the employer and with high benefit to both parties (e.g., Argyris, 1957; McGregor, 1960). The latter two positions allow the possibility of a joint employer-employee search for workplace arrangements that are mutually enhancing, while the first position implies, at best, a bargained tradeoff of costs and benefits and, at worst, a coerced solution for one party or the other.

It is known that each of the three positions is valid in some instances but not in others. A needed focus of inquiry, therefore, is to discover the circumstances under which the motivational strategies favorable to both parties are most likely to succeed. Some of these circumstances are suggested by prior work -- workforce characteristics, individual differences, local administrative and decision-making practices, and the nature of the work technology (e.g., Turner & Lawrence, 1965; Hackman & Lawler, 1971; House, 1971; Lawler, 1971; Cammann, 1976; Oldham, Hackman, & Pierce, 1976). It seems likely that the outcomes associated with alternative motivational strategies will be contingent upon such factors. They should, therefore, be taken into account in any effort to explore the feasibility of choosing motivational strategies which are of mutual rather than unilateral benefit.

A last introductory comment relates to the issues of social policy that may be accentuated or allayed by the results of inquiries such as this. If it should emerge that high levels of work motivation can be induced by means which also provide an increment of employee benefit, then an issue for local collaborative solution replaces an issue for coercion and bargaining. An opposite outcome from this inquiry would still be helpful to the extent that the nature of the choices to be made are clarified and the tradeoffs are evaluated.

Alternative Motivational Strategies

The three strategic approaches to employee motivation treated in this study correspond to the three theoretical-ideological views briefly mentioned. They are conceptually defined and further elaborated below, and operationally defined by the measurement methods to be described later.

Contingent extrinsic rewards. Operationally, this strategy involves providing valued benefits to the employee in exchange for work performance. The amount and extent of the rewards are contingent upon the amount and/or quality of the performance. The common external rewards -- i.e., rewards not derived by the employee directly from work performance, but administered by others -- include pay, promotion, job security, insurances, vacation time, and others having an economic or opportunity base. The classic and pure form of this strategy is the individual incentive pay plan which provides a fixed unit of pay increment for each unit of work performed. Modified versions provide punishments, such as discharge for excessively substandard performance, or somewhat ambiguous contingencies such as the possibility of a promotion or pay increase for superior performance. The logic of this strategy is impeccable if one assumes that better performance is costly or onerous to the worker and will ordinarily be obtained only through increments in compensating benefits. There is evidence that this motivational strategy "works", at least for some people some of the time. It is less clear, however, what the side effects may be with respect to satisfaction and mental health, and to the situational factors that moderate the main and side effects (Vroom, 1964; Lawler, 1971, 1973).

Contingent intrinsic rewards. This strategy presumes that the performance of work can be directly rewarding to the worker to the extent that the work and its environment are designed so that increments in the amount and quality of work performance are associated with increased higher-order need satisfaction. In operation, this strategy requires the provision of conditions such that superior performance enhances self-esteem, pleasurable use of abilities, needs for growth and recognition, feelings of success in dealing with suitably difficult and challenging problems, and the like. On the other hand, inferior performance would be associated with demeaning experiences, reduced opportunities for growth and recognition, etc. It is known that the introduction of such a strategy of contingent intrinsic rewards can have motivating effects, at least when superimposed upon reasonably adequate provision of non-contingent extrinsic rewards, and that it is not equally effective for all individuals and for all situations (Hackman & Lawler, 1971; Hackman & Oldham, 1974, 1975; Oldham, Hackman, & Pierce, 1976).

Rules control. This strategy, which involves control of performance through job simplification, standardization and clarification of role requirements, is based upon traditional scientific management principles as they apply to job design. Gains in work performance are sought by simplifying jobs, specifying acceptable standards of amount and quality, introducing mechanical or other non-voluntary controls on work pace, specifying the manner in which work is to be done, clarifying and narrowing the range of options left to the worker, and the like. Performance below standard is unacceptable, and above standard performance is neither expected nor routinely rewarded. The assumptions here are that the employer values a uniform, predictable and adequate level of performance, while the employee values having a job that is (relatively) undemanding, free of uncertainties, free of unnecessary effort, and clear as to methods and standards; the exchange involves equitable pay and benefits for equitable and specified work performance. The evidence, so far, suggests that this strategy can in some circumstances induce effective performance and that it may have either positive or negative side effects as to worker satisfaction (Worthy, 1950; Walker & Guest, 1952; Goldthorpe, et al., 1968, Fein, 1976).

While these three strategies can readily be distinguished they are not very well differentiated in practice. A given employer may utilize all three, applying them selectively to different classes of jobs and of people. In addition, a given job may be subject to elements of all three strategies. Few employers have an explicit, considered policy for using one over the others. For these reasons, the assumption is made in this study that the applicable strategy, or relative emphasis among the strategies, is job-specific. This requires that the strategies be measured not by reference to the employer's intention but by reference to the job occupant's perceptions and beliefs about his job and about the reward and control systems as he experiences them.

Employee Responses to Motivation Strategies

In examining the impact of these different organizational strategies for motivation and control, five different aspects of employees' responses were considered. Employee self reports of effort and involvement were chosen to represent the employee responses of particular interest to organizational leaders on the assumption that managers would rather have employees who work hard and who are involved in their jobs. In addition, three responses were examined which relate more closely to the interests of the individual employees -- work-related depression, intrinsic job satisfaction, and extrinsic reward satisfaction. Work-related depression refers to the extent that individuals report depressive symptoms in relation to their work. The measure used has been shown to be related to other negative outcomes such as poor physical health and alcoholism (Quinn & Shepard, 1974). Intrinsic and extrinsic satisfaction refer to employees' self report of feelings about their jobs and the organizational rewards such as pay and job security that they receive.

Moderating Variables

Employee responses to their environmental conditions are often found to be moderated by factors associated with the main variables under consideration. The data available for this inquiry allowed some choice among potential moderators. To explore for the presence and strength of moderating effects, eight variables were chosen --four personal characteristics, three attributes of the jobs, and one variable representing an aspect of fit between the person and his job.

Personal characteristics examined for their moderating effects include two measures of employees' needs, as well as race and sex. Responses to different organizational motivational strategies can be expected to vary as a function of the importance attributed by individuals to extrinsic and intrinsic rewards (e.g., Alderfer, 1972; Lawler, 1973; Hackman & Oldham, 1976). Race and sex were chosen as possible moderators because they are associated with differences in socialization experiences and cultural norms. People who differ on these characteristics might be expected to develop different patterns of responding to situational stimuli, and as a result, to respond differently to motivation and control strategies.

Job characteristics. Three characteristics of jobs were also examined to determine if they moderated the employees' responses to different motivational strategies. First, job autonomy was examined because it seemed possible that employees might respond differently depending on the amount of personal freedom they have in their jobs. For example, employees might be expected to respond more favorably to the intrinsic motivation strategy if they have enough autonomy to determine their own work pace or they are more directly responsible for determining their own method of performance. Similarly, such employees might respond less favorably to the strategy of using rules control if the rules impede the satisfying use of their job freedom. Second, the effects of the adequacy of resources available to complete the job were examined. Without adequate resources a job can not be done well regardless of the motivational or control strategy employed, and employees experiencing such conditions could be expected to be less involved in their jobs, less satisfied, and more depressed. Finally, the potential moderating effect of job interdependence was examined because it seemed possible that employees who must depend on other people to get their work done might respond differently to different motivation strategies than would employees who worked more independently (Thompson, 1967).

Job fit. A variable representing the degree of fit between the employee's skill level and the skill level required on the job was included as a potential moderator because there is a considerable amount of evidence indicating that skill adequacy can have an important effect on the way employees respond to their jobs, both psychologically and in terms of their performance (e.g., Caplan, et al., 1975).

Data Sources

The data were collected as part of the Effectiveness in Work Roles Study. The overall study design is described in Chapters 1 and 2. Data from the Phase I interviews were used to conduct an initial analysis of the comparative effects of different motivational strategies and of the importance of the potentially moderating conditions. The data from the Phase II interviews were used to test the replicability of the Phase I results and to note the effects, if any, from using improved measures.

The measures used for both sets of analyses were drawn from the interview portion of the data collection. They are summarized in Table 1. This table includes the variable names, sample component items where appropriate, the number of items in each measure and estimates of their internal consistency reliabilities. As can be seen from this table the measures used in Phase II were, in many cases, improved over those in Phase I by the addition or deletion of component items in order to increase the internal consistency of the multiple-item measures.

The population for Phase I consisted of 651 employees from five organizations -- a hospital, a printing company, a research and development

Table 1

Measures Used in this Report

Measures and Illustrative Components		Phase I		Phase II	
		Number of items in scale	Cronbach's Alpha	Number of items in scale	Cronbach's Alpha
Dependent:					
Effort	"Working my hardest . . . Not working hard."	1	NA	1	NA
		(7 point semantic differential)			
Job Involvement	"How involved do you feel in your job?"	3	.455	5	.653
Depression	"I feel hopeful about the future."	10	.752	10	.780
Intrinsic Satisfaction	"How satisfied are you with....... how interesting your work is?"	3	.826	4	.854
Extrinsic Satisfaction	"How satisfied are you with....... your pay?"	3	.677	3	.662
Independent:					
Contingent Reward Allocation	"If you do your job well, how likely is it....... you will be offered a better job at the place where you work?"	5	.702	4	.735

(continued.....................)

Table 1 continued

Measures Used in this Report

Measures and Illustrative Components		Phase I Number of items in scale	Phase I Cronbach's Alpha	Phase II Number of items in scale	Phase II Cronbach's Alpha
Independent continued:					
Intrinsic Motivation	"If you do your job well, how likely is it....... you will get a good feeling that you have done something important and useful?"	1	NA	2	.781
Role Clarity[1]	"I am clear about what others expect of me on my job."	3	.545	3	.606
Challenge[1]	"My job requires a high level of skill."	7	.856	6	.850
Moderators:					
Autonomy	"My job allows me to make a lot of decisions on my own."	4	.729	4	.830
Resources	"I am given enough time to do what others expect of me."	4	.606	10	.770
Interdependence	"My job is pretty much of a one person job -- there is little need for checking or meeting with others."	1	NA	1	NA

(continued.......................)

Table 1 continued

Measures Used in this Report

Measures and Illustrative Components		Phase I		Phase II	
		Number of items in scale	Cronbach's Alpha	Number of items in scale	Cronbach's Alpha
Moderators: continued:					
Extrinsic Importance	"How desirable is it....... that your pay is good?"	4	.575	3	.643
Intrinsic Importance	"How desirable is it....... that you get a sense of accomplishment from your work?"	5	.776	7	.804
Race		1	NA	1	NA
Sex		1	NA	1	NA
Skillfit[2]		1	NA	1	NA

---

[1] The Role Clarity and Challenge (reversed) scales were multiplied together to form the rules control scale.

[2] The Skillfit variable was formed by subtracting the actual education level from the level needed for the job (self report).

laboratory, and two plants that manufacture automobile accessories. The Phase II population consisted of 272 employees from the hospital and the two plants, most of whom were members also of the Phase I population. Aside from the smaller number of respondents in Phase II, there were no major compositional differences in the two populations. Though not representative of any defined population base, the respondents do represent a variety of common occupations within common kinds of work establishments. The populations include both supervisory and nonsupervisory employees, but exclude those working fewer than 20 hours per week and those new in their jobs. Compared to the national population of employed adults, these populations contain a somewhat elevated proportion of blacks, women, single people, and young people. Fuller descriptions of the populations appear in Chapters 1 and 2 of this report.

Analysis Plan

The analysis was conducted in two main stages. The first was intended to test whether the relationship between each of the motivational strategies and the dependent measures of employee responses were significantly different depending on the level of the moderating variable. The second stage was intended to apply this information by assessing the impact of the alternative motivational strategies on employees' response for the population as a whole and for subpopulations selected to remove the confounding effects of the moderators.

The interaction analyses were carried out in four steps. First, all possible first-order terms were formed by multiplying the dependent measures times the moderator variables for the Phase I data. The test for the significance of these interaction terms is analogous to the one used in testing for moderator effects in a regression model (Kerlinger & Pedhauser, 1973). The only difference between the standard test for interaction and the test used here was that multiple dependent variables were involved and canonical analyses were used instead of regression analyses (see Cammann, 1976, for another example of this type of analysis). A Chi Square Test based on Wilks' Lambda was used to test the significance of interactions (see Tatsuoka, 1971, for a description of this procedure). This analysis showed which moderating conditions, singly, had a significant effect.

The second step involved examining the second order interactions of the motivational strategies with two moderating conditions at a time. This was necessary in order to examine the possibility that the moderating conditions themselves interact in moderating employee responses to the different motivation strategies. Third and higher order interactions were not examined both because they would be difficult to meaningfully interpret, and because the sample was not large enough to allow a full exploration of their implications.

The third step involved adding all of the significant moderators as main effects, and then their interaction terms, into a set of canonical analyses. The purpose of this analysis was to make sure that each interaction was contributing significant additional explanatory power above that which could already be explained by the main effects. Since the interaction analysis had been conducted separately it was possible that they would not.

The final step in the interaction analysis involved the analysis of the Phase II data. Step 3 was repeated using the Phase II data to test the replicability of the interaction analyses. Moderating conditions which did not replicate were not analyzed further unless there was some characteristic of the Phase II population which might have caused the lack of replication.

The second stage of analysis involved examining the relationships between the organizational strategies for motivation and control and employee responses for the total Phase I and Phase II populations and for each of the subsamples identified by significant moderators. The subpopulations were formed by splitting the total population into two or three groups on each of the moderators to allow examination of the relationships within different levels of the moderating variables.

Results

Prior to the two major stages of the analysis, the correlations among the independent and moderating variables were examined. These correlations are shown in Table 2. As can be seen in this table, the measures tend to be only moderately intercorrelated. There is a distinct patterning of signs revealing a tendency for the organizations in this study to use the contingent extrinsic reward strategy and the intrinsic motivation strategy together, and not use these strategies in conjunction with rules control. The strategies of contingent external rewards and intrinsic motivation tend to be used for jobs that are high on autonomy and availability of resources. and that are occupied by individuals who value intrinsic rewards and who are underqualified for their jobs. The rules control strategy tends to be used in the opposite situation. This multi-collinearity indicates that the motivational strategies these organizations used were not independent of the moderating variables. This may potentially

Table 2

Correlation of Independent and Moderating Variables

	Phase I			Phase II		
	Independent Variables					
Variables	Contingent Reward Allocation	Intrinsic Motivation	Rules Control	Contingent Reward Allocation	Intrinsic Motivation	Rules Control
Moderator:						
Autonomy	.34**	.31**	-.35**	.42**	.43**	-.52**
Resources	.22**	.17**	.21**	.30**	.31**	-.09
Interdependence	.02	-.03	.15**	.18**	.21**	-.21**
Extrinsic Importance	.03	.12**	-.02	.08	.18**	.03
Intrinsic Importance	.16**	.27**	-.25**	.18**	.30**	-.32**
Race[a]	-.06	-.07	.14**	.05	-.23**	.14*
Sex[b]	-.17**	-.07	.17**	-.10	-.02	.20**
Skillfit	-.26**	-.30**	.41**	-.33**	-.29**	.43**

(continued . . . .)

Table 2 continued

Correlation of Independent and Moderating Variables

	Phase I			Phase II		
	Independent Variables					
Variables	Contingent Reward Allocation	Intrinsic Motivation	Rules Control	Contingent Reward Allocation	Intrinsic Motivation	Rules Control
Independent:						
Contingent Reward Allocation						
Intrinsic Motivation	.32**			.40**		
Rules Control	-.27**	-.29**		-.22**	-.30**	

*p < .05
**p < .01

[a] Dichotomous variable: 1 = White; 2 = Other.

[b] Dichotomous Variable: 1 = Male; 2 = Female.

attenuate the results of the analyses since it may lead to restriction of the range of the independent variables within levels of the moderators.

The Moderator Analyses: Stage I Results

The results of the Phase I moderator analyses are summarized in Table 3. This table shows the significance of the Chi Square values for the canonical analyses with just the main effects included (each motivation strategy and each moderating condition), and also the Chi Square values for the canonical analyses with the interaction terms as well. The significance levels of the interactions were determined by testing the significance of the differences in Chi Squares (df = 5, in the test).

For the contingent extrinsic reward strategy, the Phase I data results indicate that two moderating variables -- race and skillfit -- had significant effects. Extrinsic needs, sex, autonomy, and resources had significant main effects but not as moderators. When each of the first order moderators was added sequentially into a canonical analysis after race and skillfit (as main effects) both race and skillfit were found to contribute additional explanatory power so they were retained. The two-way interactions were next entered into the canonical analyses. Only those variables which had proven to have significant first order interactions with the motivating strategies were used in this stage of the analysis. The extrinsic by skillfit by race interaction was found to be significant.

For the rules control strategy, race, sex, job interdependence, and skillfit were significant moderators. The remaining job characteristics and individual needs were not. When each of the first order moderators

Table 3

Significance Levels for Chi Squared Statistics Used in Moderator Analysis

Motivational Strategy	Contingent Reward Allocation		ules Control		Intrinsic Motivation	
Moderators	Main Effects[a]	Moderator Effect[b]	Main Effects[a]	Moderator Effect[b]	Main Effects[a]	Moderator Effect[b]
First Order:						
Extrinsic needs	.05	NS	Not Tested		Not Tested	
Intrinsic needs	Not Tested		.05	NS	.05	NS
Race	.05	.05[e]	.05	.05[d]	.05	.05[e]
Sex	.05	NS	NS	.05[d]	.05	NS
Job autonomy	.05	NS	.05	NS	.05	.05[e]
Resources	.05	NS	.05	NS	.05	NS
Interdependence	NS	NS	NS	.05[c]	NS	NS
Skillfit	.05	.05[d]	.05	.05[c]	.05	.05[c]

(continued . . . .

Table 3 continued

Significance Levels for Chi Squared Statistics Used in Moderator Analysis

Motivational Strategy	Contingent Reward Allocation		Rules Control		Intrinsic Motivation	
Moderators	Main Effects[a]	Moderator Effect[b]	Main Effects[a]	Moderator Effect[b]	Main Effects[a]	Moderator Effect[b]
Second Order:						
Skillfit x Race		$.05^d$		NS		$.05^e$
Skillfit x Sex	Not Tested			$.05^e$	Not Tested	
Race x Sex	Not Tested			NS	Not Tested	
Autonomy x Skillfit	Not Tested		Not Tested			NS
Autonomy x Race	Not Tested		Not Tested			NS

[a] Main effects significance test reflects the significance of the moderator as a linear main effect.

[b] Moderator effect significance test reflects the significance of the moderator and independent variable interaction.

[c] Although these interactions were significant when tested alone, they did not contribute additional explanatory power when added sequentially into the canonical analysis after the main effects.

[d] Interactions which were replicated in the Phase II data.

[e] Interactions which were not replicated in the Phase II data.

was added sequentially into a canonical analysis, after race, sex, job interdependence, and skillfit (as main effects) the rules control by interdependence and the rules control by skillfit interactions were eliminated, while the other two interactions were retained. In testing the second order interactions, only the rules by skillfit by sex interaction was found to be significant.

For the intrinsic motivation strategy the significant moderators were race, autonomy, and skillfit. The relationship between employee's responses and the intrinsic motivation strategy were not moderated by individual need strength, sex, resources, or job interdependence. When each of the moderators was added into a canonical analysis after race, autonomy, and skillfit the intrinsic by skillfit interaction dropped out while the other two first order moderators were retained. In testing the higher order interactions only the intrinsic by skillfit by race interaction was found to be significant.

The next analysis step was to replicate the foregoing moderator analyses using Phase II data and limiting the analysis to those variables which had proved significant in the Phase I data. In the analysis involving the measure of the contingent extrinsic reward strategy the moderating effects of skillfit and skillfit by race replicated, while the effects of race did not. The replication analysis for the rules control strategy indicated that the moderating effects of race and sex replicated but the moderating effect of skillfit by sex did not. The results for the intrinsic motivation strategy indicated that none of the moderating effects were replicated.

Before concluding that the moderators which did not replicate did not really have a significant effect on the nature of the relationships between the strategies for motivation and employee responses, the population characteristics for Phase I and Phase II were examined to see if differences might explain the observed results. In the case of the race moderating variable (which did not replicate for the contingent reward allocation and the intrinsic motivation analyses) the differences in the two populations might explain the lack of replication. The number of non-whites in the sample went from 127 to 71. Thus, it is possible that the failure of replication for race might reflect the decreased stability of relational estimates due to the lower absolute number of non-white respondents in the Phase II population. For this reason it was decided to retain race as a moderator even though it had not replicated. On the same grounds it was decided to explore the race by skillfit interaction for the intrinsic motivation strategy.

No plausible reason could be identified for the failure to replicate the skillfit by sex interaction as a moderator of the relationship between the measure of rules control and the measure of employee responses. The samples and the distribution of scores for both Phase I and Phase II populations were similar. It was concluded that the significant finding in Phase I was probably a statistical artifact.

With regard to the lack of replication of the intrinsic by autonomy interaction, a sampling problem appears to be present. While 60% of the Phase I sample reported high autonomy, only 48% of the Phase II sample did so. Because this lack of replication may have been due to a significant change in our sample the intrinsic by autonomy interaction was retained for subsequent analytic steps.

To summarize: The results of the moderator analyses indicated that the relationships between the contingent extrinsic reward strategy and employee responses should be explored for subsamples based on race, skillfit, and the race by skillfit interaction; The relationships between rules control and employee responses should be explored for subsamples based on sex and race; and, The relationships between intrinsic motivation and employee responses should be examined for subsamples based on autonomy, race, and race by skillfit interactions. Employee needs, job interdependence, and resources did not appear to moderate any of the relevant relationships and were, therefore, not used to divide the population.

Results: Analysis Stage II

The next stage of the analysis involved examining the relationships between the three alternative motivation strategies and the five employee responses for the total population (both Phase I and Phase II) and for the relevant subsamples based upon variables identified by the moderator analyses. The first step was to split the total population into subsamples based on the moderator variables. This was accomplished in different ways for different moderators. For the sex and race variables dividing was dichotomous: Male and female; White and non-white. For the skillfit measures the population was divided into three groups: The employees who said their jobs required more education than they had; The same; or, Less. For autonomy the population was divided on the total distribution of autonomy scores which showed a clean dichotomous break point.

Results for the contingent extrinsic reward strategy. The relationships between the measure of contingent extrinsic reward allocation and the measures of employee responses are shown in Table 4b for the total population and for the relevant subsamples for each of the two phases of data collection. The mean scores for all of the variables for each subsample are shown in Table 4a. These results indicate that, in general, contingent extrinsic reward allocation had a slight but positive relationship with self-reports of effort and a stronger positive relationship with involvement. It was also generally positively related to both intrinsic and extrinsic satisfaction and negatively related to depression. The moderator analyses (subsamples) produced few consistent differences from this basic relational pattern. Whites differed from non-whites in the strength of relationships between the perception of contingent reward allocation and involvement in the Phase I population, but this finding was not replicated in Phase II. Underqualified whites showed a stronger positive relationship between the contingent extrinsic reward allocation measure and effort than did qualified and overqualified whites or non-whites, but, again, this finding was not replicated in Phase II. One of the few clearly replicated results was the generally low level of relationship of contingent extrinsic reward allocation to employee responses for qualified whites, i.e., those with a "fit" of qualification to job requirements. The canonical correlations for this sample (r's = .38; .38) were the lowest in both phases of the study (next lowest, r's = .45; .51; .55; .60) indicating that this group does not appear to respond as strongly to the contingent extrinsic reward allocation strategy as do the other groups.

These results lead to the observation that the moderating conditions had an operable effect on the strength of the relationships between contingent extrinsic reward allocation and employee responses, but had little effect on the direction of these relationships. Also, in general, the strengths of the moderating effects were not sufficient to replicate consistently.

Table 4a

Mean Response to Contingent Reward Allocation Strategy for Total Population and Relevant Subpopulations

Phase	Total Population	White	Non-White	Under-Qualified	Over-Qualified	Over-Qualified	Under-Qualified White[1]	Qualified White	Over-Qualified White	Qualified Non-White	Over-Qualified Non-White
Phase I (N)	(651)	(513)	(127)	(105)	(284)	(243)	(92)	(233)	(178)	(48)	(63)
Contingent Reward Allocation	2.12	2.15°	2.03	2.43°	2.16°	1.95	2.41°	2.18	1.96	2.06	1.92
Effort	6.00	6.10°	5.56	6.38°	6.11	5.69	6.32°	6.17	5.88	5.78	5.15
Involvement	3.30	3.35°	3.11	3.56°	3.39	3.08	3.56°	3.41	3.15	3.27	2.89
Depression	2.01	2.00	2.06	1.93°	1.97	2.10	1.95°	1.98	2.05	1.95	2.18
Intrinsic Satisfaction	2.94	2.94	2.93	3.28°	3.15	2.54	3.26°	3.24	2.50	3.15	2.66
Extrinsic Satisfaction	3.03	3.10°	2.76	3.15°	3.06	2.93	3.20°	3.09	3.04	2.93	2.61
Phase II (N)	(266)	(195)	(71)	(54)	(115)	(88)	(45)	(80)	(60)	(35)	(28)
Contingent Reward Allocation	2.01	1.99	2.09	2.31°	2.11	1.73	2.31	2.08	1.66	2.20	1.87
Effort	5.89	5.88	5.93	5.91	6.07	5.65	5.84	6.06	5.65	6.09	5.64
Involvement	2.70	2.70	2.71	2.99°	2.74	2.44	3.03°	2.72	2.38	2.79	2.56
Depression	1.88	1.85	1.97	1.71°	1.81	2.08	1.73°	1.81	2.04	1.83	2.16
Intrinsic Satisfaction	3.04	3.05	3.03	3.32°	3.18	2.64	3.37°	3.22	2.52	3.11	2.89
Extrinsic Satisfaction	3.12	3.19	2.95	3.24°	3.23	2.87	3.32°	3.29	2.90	3.07	2.81

°The means for the split are significantly different ($p < .01$).

[1] Under-qualified nonwhites are excluded from this table because of the small sample size (N < 8 in both phases).

Table 4b

Correlations between Contingent Reward Allocation Strategy and Employee Responses for Total Population and Relevant Subpopulations

Phase	Total Population	White	Non-White	Under-Qualified	Qualified	Over-Qualified	Under-Qualified White[1]	Qualified White	Over-Qualified White	Qualified Non-White	Over-Qualified Non-White
Phase I (N)	(651)	(513)	(127)	(105)	(284)	(243)	(92)	(233)	(178)	(48)	(63)
Effort	.13**	.10*	.19*	.22*	.03	.13*	.22*	-.02	.13	.14	.15
Involvement	.27**	.33**	.04	.33**	.11	.30**	.38**	.14*	.44**	-.13	-.01
Depression	-.31**	-.31**	-.31**	-.16	-.29**	-.31**	-.22*	-.17**	-.33**	-.44**	-.24
Intrinsic Satisfaction	.38**	.38**	.40**	.24*	.28**	.41**	.26*	.28**	.40**	.36*	.40**
Extrinsic Satisfaction	.29**	.27**	.32**	.30**	.28**	.26**	.30**	.26**	.25**	.31*	.28*
Phase II (N)	(266)	(195)	(71)	(54)	(115)	(88)	(45)	(80)	(60)	(35)	(28)
Effort	.08	.16*	-.09	.16	-.09	.09	.19	-.18	.33**	.05	-.38*
Involvement	.35**	.31**	.47**	.21	.28**	.33**	.15	.09	.38**	.69**	.21
Depression	-.25**	-.32**	-.15	-.26	-.05	-.33**	-.47**	-.04	-.43**	-.07	-.22
Intrinsic Satisfaction	.44**	.44**	.43**	.34*	.22*	.54**	.43**	.10	.56**	.46**	.48**
Extrinsic Satisfaction	.35**	.37**	.32**	.26	.33**	.31**	.26	.27*	.42**	.51**	.16

*p < .05
**p < .01

[1] Under-qualified nonwhites are excluded from this table because of the small sample size (N < 8 in both phases).

Results for the rules control strategy. The results of the analyses for rules control are shown in Tables 5a and 5b. They indicate that the rules control measure was slightly and negatively related to the measure of effort, and more strongly negatively related to the measures of involvement and intrinsic satisfaction. It was also slightly positively related to the measure of work related depression and, generally, unrelated to the measures of extrinsic satisfaction. The subsample analyses indicated that non-whites and women were slightly more likely to have jobs which they characterized as being high on rules control and their job responses were, generally, less strongly related to the rules control measure. Again, as was the case for the contingent extrinsic reward allocation results, the moderating variables appeared to effect the strength of the relationships between the rules control strategy and employee responses but not the direction of the relationships.

Results for the intrinsic motivation strategy. The results of the analyses involving the intrinsic motivation strategy are shown in Tables 6a and 6b. In the Phase I sample the measure of the intrinsic motivation strategy was moderately positively related to the measures of effort, involvement, intrinsic satisfaction, and extrinsic satisfaction and moderately negatively related to the measure of work-related depression. The subsample analyses replicated that pattern of results in all cases but the relationships, generally, were weaker for the underqualified whites and qualified non-whites (although both of these groups showed a higher association between intrinsic motivation and effort). The association also tended to be weaker for employees with high autonomy than for employees with low autonomy.

The results for the Phase II popluation were similar. The relationships

Table 5a

Mean Responses to Rules Control Strategy for Total Population and Relevant Subpopulations

Population	Total Sample	White	Non-White	Male	Female
Phase I (N)	(647)	(513)	(127)	(330)	(318)
Rules Control	6.48	6.29°	7.26	6.02°	6.96
Effort	6.00	6.10°	5.56	6.05	5.94
Involvement	3.30	3.35°	3.11	3.39°	3.21
Depression	2.01	2.00	2.06	1.98	2.04
Intrinsic Satisfaction	2.94	2.94	2.93	2.95	2.92
Extrinsic Satisfaction	3.03	3.10°	2.76	3.04	3.03
Phase II (N)	(266)	(195)	(71)	(142)	(115)
Rules Control	7.25	6.94°	8.03	6.80	7.81
Effort	5.88	5.88	5.93	5.72°	6.11
Involvement	2.70	2.70	2.71	2.82°	2.54
Depression	1.88	1.85	1.97	1.85	1.92
Intrinsic Satisfaction	3.04	3.05	3.03	3.07	3.01
Extrinsic Satisfaction	3.12	3.19°	2.95	3.11	3.14

°The means for this split are significantly different ($p < .01$).

Table 5b

Correlations between Rules Control Strategy and Employee Responses for Total Population and Relevant Subpopulations

Populations	Total Sample	White	Non-White	Male	Female
Phase I (N)	(647)	(513)	(127)	(330)	(318)
Effort	-.16**	-.13**	-.17	-.17**	-.15**
Involvement	-.39**	-.43**	-.19*	-.41**	-.33**
Depression	.18**	.20**	.03	.13*	.21**
Intrinsic Satisfaction	-.36**	-.43**	-.10	-.43**	-.29**
Extrinsic Satisfaction	-.04	.00	-.08	-.08	-.01
Phase II (N)	(266)	(195)	(71)	(142)	(115)
Effort	-.12	-.08	-.23	-.19*	-.09
Involvement	-.47**	-.52**	-.35**	-.49**	-.41**
Depression	.23**	.25**	.13	-.30**	.14
Intrinsic Satisfaction	-.36**	-.47**	-.05	-.49**	-.20*
Extrinsic Satisfaction	-.17**	-.16*	-.12	-.33**	-.02

*p < .05
**p < .01

Table 6a

Mean Responses to Intrinsic Reward Contingency for Total Population and Relevant Subpopulations

Phase	Total Population	White	Non-White	Low Autonomy	High Autonomy	Under-Qualified White[1]	Qualified White	Over-Qualified White	Qualified Non-White	Over-Qualified Non-White
Phase I (N)	(651)	(513)	(127)	(261)	(387)	(92)	(233)	(178)	(48)	(63)
Intrinsic Contingency	3.08	3.11	2.92	2.77	3.29	3.45	3.50	2.78	2.95	2.80
Effort	6.00	6.10	5.56	5.85	6.10	6.32	6.17	5.88	5.78	5.15
Involvement	3.30	3.35	3.11	3.10	3.44	3.56	3.41	3.15	3.27	2.89
Depression	2.01	2.00	2.06	2.13	1.93	1.95	1.98	2.06	1.95	2.18
Intrinsic Satisfaction	2.94	2.94	2.92	2.52	3.22	3.26	3.14	2.50	3.15	2.66
Extrinsic Satisfaction	3.03	3.10	2.76	2.78	3.19	3.20	3.09	3.04	2.93	2.61
Phase II (N)	(266)	(195)	(71)	(137)	(124)	(45)	(80)	(60)	(35)	(28)
Intrinsic Contingency	3.33	3.43	3.05	3.05	3.64	3.64	3.46	3.18	3.22	2.80
Effort	5.88	5.88	5.93	5.80	5.94	5.84	6.06	5.65	6.09	5.64
Involvement	2.70	2.70	2.71	2.49	2.93	3.03	2.72	2.38	2.79	2.56
Depression	1.88	1.85	1.97	2.02	1.73	1.73	1.81	2.04	1.83	2.16
Intrinsic Satisfaction	3.04	3.05	3.03	2.72	3.39	3.37	3.22	2.52	3.11	2.89
Extrinsic Satisfaction	3.12	3.19	2.95	2.88	3.39	3.32	3.24	2.90	3.07	2.81

[1] Under-qualified nonwhites are excluded from this table because of the small sample size (N < 8 in both phases).

Table 6b

Correlations between Intrinsic Contingency Strategy and Employee Responses for Total Population and Relevant Subpopulations

Phase	Total Population	White	Non-White	Low Autonomy	High Autonomy	Under-Qualified White[1]	Qualified White	Qualified White	Qualified Non-White	Over-Qualified Non-White
Phase I (N)	(651)	(513)	(127)	(261)	(387)	(92)	(233)	(178)	(48)	(63)
Effort	.22**	.20**	.25**	.24**	.19**	.21*	.16*	.16*	.33*	.14
Involvement	.34**	.34**	.33**	.32**	.28**	.15	.31**	.33**	.14	.36**
Depression	-.35**	-.37**	-.29**	-.41**	-.21**	-.12	-.32**	-.44**	-.22	-.29*
Intrinsic Satisfaction	.41**	.44**	.27**	.39**	.30**	.10	.37**	.49**	.26	.27*
Extrinsic Satisfaction	.21**	.18**	.26**	.13*	.18**	-.01	.20**	.23**	.39**	.16
Phase II (N)	(266)	(195)	(71)	(137)	(124)	(45)	(80)	(60)	(35)	(28)
Effort	.18**	.25**	.07	.24**	.14	.08	.10	.35**	.43**	-.30
Involvement	.31**	.39**	.20*	.21**	.23*	.04	.32**	.41**	.41*	-.14
Depression	-.40**	-.44**	-.29*	-.36**	-.28**	-.39**	-.32**	-.44**	-.16	-.31
Intrinsic Satisfaction	.44**	.51**	.34**	.36**	.25**	.41**	.47**	.40**	.48**	.19
Extrinsic Satisfaction	.35**	.33**	.30*	.32**	.11	.31*	.14	.36*	.48**	.12

*p < .05
**p < .01

[1] Under-qualified nonwhites are excluded from this table because of the small sample size (N < 8 in both phases).

for the total population showed no major differences. The results for the white subsample were generally stronger and those for the non-whites generally weaker. The results for the high and low autonomy subsamples were again similar with the high autonomy subsample showing slightly weaker relationships. The results for underqualified whites and for qualified non-whites were not weaker for the Phase II sample as they had been in the Phase I sample, and the results for overqualified whites showed non-significant changes in the direction of some of the relationships.

Overall, the general pattern of results indicated for both Phase I and Phase II populations that the relationships between the alternative motivation strategies and the employee responses were consistent across subpopulations and that the major effects of the moderating variables were in the strength of the relationships, not in their direction.

Discussion and Interpretation

The introduction to this report stated a number of issues to which the inquiry was addressed. These were concerned, generally, with issues of the compatibility among alternative employer strategies for inducing employee effort and involvement; whether a given strategy or combination of strategies tends to favor outcomes of interest to employers or to employees; and whether there are strong moderating effects that suggest a need to tailor the motivational strategy to local characteristics of the workforce or jobs involved. While the inquiry lacks sufficient scope and measurement precision to provide definitive answers to such questions, the results appear to add significantly to the accumulating evidence from which definitive answers will ultimately emerge.

Compatibility of alternative motivational approaches. Three conceptually distinct approaches were treated. They emphasized extrinsic rewards, intrinsic rewards, and rule-and-role specification as inducements to high work effort and job involvement. The results of the inquiry are unequivocal in supporting the view that the first two approaches are compatible while the third is not. By compatibility we mean here both the feasibility of using two or more of the approaches simultaneously in a given work organization and also compatibility in the sense that the outcomes of the jointly-used approaches are similar, and not contradictory. The evidence shows that the contingent extrinsic rewards and the contingent intrinsic rewards strategies tend to be used together in the organizations studied. It is apparently not necessary to choose between them as incompatible alternatives, at least not in the organizational situation studied here (Staw, 1975). The third strategy involving motivation and control through job simplification, specification of rules, and impersonal controls appears to be incompatible with the others in two respects: (1) Where the latter is used, the others tend not to be used; and (2) It's association with criterion variables suggest that the outcomes are opposite — at least with reference to the criteria used in this inquiry.

Outcome congruence and tradeoffs. The contingent extrinsic reward and contingent intrinsic reward strategies display criterion congruity in the sense that neither shows any tendency to be linked with outcomes that favor one of the parties of interest over the others. That is, the evidence supports the conclusion that both strategies are associated with advantages to the employer in the form of relatively high work effort and work involvement, and also advantages to the employees in the form of need satisfaction and moderation of depressive symptoms. The situation is not one in which one party is necessarily advantaged at the expense of the other. The same observation about congruence and tradeoffs can be made with respect to the clear rules and roles strategy except that in this case the associations are reversed so that the two parties of interest are both similarly disadvantaged.

Moderating effects. There is very little support in this inquiry for the notion that optimum motivational strategies are substantially different for situations where distinctly different employee populations or job attributes are found. On the contrary, there is evidence that the moderating factors explored have: (1) Only weak effects; and (2) Have these effects similarly across a variety of contrasting situations. That is, the main conclusions of this inquiry appear to apply across several contrasting populations and jobs even when these are chosen to maximize the power of the hypothesized moderating factors. This is not to say that some common array of motivational conditions is universally and equally effective everywhere, but only that the local optimizing is likely to be in the form of "fine tuning" rather than switching to some contrasting approach or set of guiding principles. It was anticipated that the prevalent signs of association between a motivational strategy and the several criteria, whether positive or negative, would be reversed for some subpopulations but this condition was not found.

At this point a comment is needed with respect to a minor supplemental analysis of the data. The discovery of the prevalence of simultaneous use by employers of both the extrinsic and intrinsic motivational strategies raised a question about the relationships between the two. Staw (1975), for example, has asserted that in non-voluntary organizations intrinsic and extrinsic motivations can operate in a complementary manner. In contrast, the experiments of Deci (1975) suggest that an individual's response to one source of motivation might diminish response to the other. The data in the present study indicate that the two are significantly independent and also additive. They are positively, but modestly, correlated in application, and their partial correlations with the various dependent variables (criteria) show that each makes a contribution independently of the other. Further, a canonical analysis indicated that the two motivational strategies did not interact significantly with each other. The evidence suggests a simple additive and complementary relationship.

While the foregoing results appear to be rather clear, to be well supported by the data and analytic procedures, and to be of obvious implication as to preferred strategies in employee motivation, some cautionary comments need to be made.

The main caution arises from the limited scope of the variables available for inclusion in this study. For example, the roster of dependent (criterion) variables, while providing some diversity and including reference to both employer and employee priority values, is far from a complete and balanced representation of the domain. An example will make this point clearly. This inquiry presents a consistently adverse view of the "clear rules and roles" motivational strategy, but this research outcome may well be a consequence of the omission of certain key variables relevant to the evaluation of that motivational strategy. One might suppose, for example, that given a long-linked, capital-intensive, mass-production, work technology the applicibility of the alternative strategies might be limited, the cost effectiveness of work might be high even though effort and involvement are low, and that there might be a valid tradeoff situation in which managerial interests conflict with the interests of the employees. The present study contained no measures of cost effectiveness nor of technological constraints. Other examples could be mentioned.

A second limitation of this inquiry lies in the cross-sectional analysis design. The data did not permit the two-phase panel measurements to be used in a more definitive causal design. The results have been described in language that suggests a direction of causation such that the motivational strategies caused the variation in dependent variables. It is possible that some reverse causation is present such that, for example, employees who exert high effort may in time be provided with a motivational approach less dependent upon the imposition of rules and explicit role requirements.

Finally, the populations of persons and jobs available for this study, while in important ways typical of prevalent workplace conditions, may not include extreme cases or combinations of conditions which, if present in larger numbers, might have modified the conclusions reached.

Quinn, R. P., & Shepard, L. J. *The 1972-73 quality of employment survey: Descriptive statistics with comparison data from the 1969-70 survey of working conditions*. Ann Arbor: Survey Research Center, 1974.

Srivastra, S. et al. *Job satisfaction and productivity*. Cleveland: Case-Western Reserve University, 1975.

Staw, B. M. Attribution of the "causes" of performance: A general alternative interpretation of cross-sectional research on organizations. *Organizational Behavior and Human Performance*, 1975, *13*, 414-432.

Tatsuoko, M. M. *Multivariate analysis: Techniques for educational and psychological research*. New York: John Wiley and Sons, Inc., 1971.

Thompson, J. D. *Organizations in action*. New York: McGraw-Hill, 1967.

Turner, A. N., & Lawrence, P. R. *Industrial jobs and the worker*. Boston: Howard University Graduate School of Business Administration, 1965.

Vroom, V. H. *Work and motivation*. New York: John Wiley, 1964

Walker, C. R., & Guest, R. H. *The man on the assembly line*. Cambridge: Harvard University Press, 1952.

Worthy, J. C. Organizational structure and employee morale. *American Sociological Review*, 1950, *15*, 169-179.

References

Alderfer, C. P. *Existence, relatedness, and growth: Human needs in organizational settings*. New York: The Free Press, 1972.

Argyris, C. *Personality and organization*. New York: Harper, 1957.

Cammann, C. C. Effects of the use of control systems. *Accounting, Organizations and Society*, 1976, *1*, 301-313.

Caplan, R. D., Cobb, S., French, J. R. P., Jr., Harrison, R. V., & Pinneau, S. R., Jr. *Job demands and worker health*. U.S. Department of Health, Education and Welfare, 1975.

Deci, E. L. *Intrinsic motivation*. New York: Plenum Press, 1975.

Fein, M. Motivation for work. In R. Dubin (Ed.), *Handbook of work, organization, and society*. Chicago: Rand McNally College Publishing Company, 1976.

Goldthorpe, J. H., Lockwood, D., Bechhofer, F., & Platt, J. *The affluent worker: Individual attitudes and behavior*. London: Cambridge University Press, 1968.

Hackman, J. R., & Lawler, E. E. III. Employee reactions to job characteristics. *Journal of Applied Psychology*, 1971, *55*, 259-286.

Hackman, J. R., & Oldham, G. R. Development of the job diagnostic survey. *Journal of Applied Psychology*, 1975, *60*, 159-170.

Hackman, J. R., & Oldham, G. R. Motivation through the design of work: Test of a theory. (Technical Report No. 6) New Haven, Conn.: Yale University, Department of Administrative Sciences, 1974.

House, R. J. A path-goal theory of leader effectiveness. *Administrative Science Quarterly*, 1971, *2*, 321-339.

Katzell, R. A., Yankelovich, D. *Work, productivity, and satisfaction*. New York: Psychological Corporation, 1975.

Kerlinger, F. N., & Pedhazur, E. J. *Multiple regression in behavioral research*. New York: Holt, Rinehart and Winston, Inc., 1973.

Lawler, E. E. III. *Motivation in work organizations*. Monterey, Ca.: Brooks/Cole Publishing Company, 1973.

Lawler, E. E. III. *Pay and organizational effectiveness: A psychological view*. New York: McGraw-Hill, 1971.

McGregor, D. *The human side of enterprise*. New York: McGraw-Hill, 1960.

Oldham, G. R., Hackman, J. R., & Pierce, J. L. Conditions under which employees respond positively to enriched work. *Journal of Applied* Psychology, 1976, *61*, 395-403.

Chapter 20

IMPACT OF PERFORMANCE-CONTINGENT REWARDS ON JOB SATISFACTION

by

Nina Gupta

ABSTRACT

This analysis confirms a direct relationship between performance-contingent rewards and satisfaction with these rewards independently of the actual reward amount. That is, the contingency of the rewards is related to satisfaction not merely because of the receipt of rewards following good performance but because the perception of the contingency is itself satisfying.

Several psychological processes are advanced as the bases for the prediction of this result and for the understanding of its meaning. These processes relate to the desire for equity, the desire for self-control of transactions with one's environment, and desire for certainty (low ambiguity) in such transactions. The results show that locus of control is a significant factor in the present study; tolerance for ambiguity is not a significant factor; perceptions of equity could not be examined.

The results suggest that the institution of performance-based variable reward systems, as compared with the allocation of rewards on categorical grounds other than performance (e.g., seniority, age, sex), will have multiple benefits for the employing organization. In addition to the motivation effects of performance-contingent rewards there may be provided an added measure of job satisfaction.

Chapter 20

IMPACT OF PERFORMANCE-CONTINGENT REWARDS ON JOB SATISFACTION*

The reward system of an organization includes some definition of the criteria by which rewards are allocated to members. An organization may provide various rewards -- e.g., pay, promotions -- uniformly to all individuals, or selectively to different subpopulations within the organization, or uniquely to each individual based on various ability and non-ability criteria. The present study examines the impact of a common criterion for reward allocation, namely, rewards provided to the individual on the basis of performance. The basic proposition to be tested in this study is that the provision of rewards that are contingent on good performance has positive effects on job satisfaction. Also examined are the moderating effects of personality and situational characteristics on this relationship.

The idea that rewarding good performance leads to sustained or improved performance has been the subject of much theory, variously labelled expectancy theory (e.g., Porter & Lawler, 1968), instrumentality theory (Graen, 1969), path-goal theory (Georgopoulos, Mahoney, & Jones, 1957), and behavior modification (Skinner, 1969). Most of the research within the context of these theories has explored the effects of performance-contingent rewards on motivation, effort, or performance; relatively little has been directed to the effects of such a reward system on the satisfaction of employees. A major criterion for assessing the effectiveness of an organizational reward system, however, is the impact such a system has on the job satisfaction of organizational members. It is important, therefore, to examine the relationship of performance-contingent rewards and satisfaction.

---

*This Chapter is abstracted from Gupta (1975). Partial support was provided by the Department of Psychology, The University of Michigan.

A distinction should be made here between intrinsic rewards, inherent in the performance of an activity, and extrinsic rewards, external to the activity although the activity is instrumental to obtaining these rewards. One of the extrinsic rewards that individuals most uniformly seek from their membership in work organizations is money. The ensuing discussion, therefore, gives special attention to the extrinsic reward of pay as well as to intrinsic rewards.

Intrinsic Rewards and Satisfaction

Intrinsic rewards are, by definition, satisfying. Positive relationships between various intrinsic rewards and overall job satisfaction have been reported by several researchers (e.g., Herzberg, Mausner, & Snyderman, 1959; Vroom, 1964). These studies, however, did not specify clearly whether the intrinsic rewards were contingent on good performance. A few studies exist where this performance contingency characteristic was explicitly spelled out (e.g., Graen, 1969; Hackman & Lawler, 1971; Mitchell & Albright, 1972; Arvey & Mussio, 1973). While the treatment of satisfaction as a criterion in these studies was limited, a positive relationship is suggested between contingent/intrinsic rewards and satisfaction.

It was assumed for this inquiry that intrinsic rewards are self-administered and that the degree of performance contingency is also self-determined, even though others may aid in supplying conditions that maximize the likelihood of either or both. Supporting evidence for these assumptions is provided later. The degree of independence between intrinsic reward level and degree of contingency is shown by their relatively low correlation ($r = .41$). That the degree of contingency does indeed vary among workers is shown by the obtained mean score of 3.08 (on a four point scale) and the standard deviation of .98.

Performance-Contingent Pay and Satisfaction

Research evidence has also emerged on the effects of performance-contingent pay rewards on satisfaction, within the framework of expectancy theory. Again, the effects on satisfaction were not examined systematically and were usually a by-product of motivation research. Among the expectancy theory studies on performance-contingent pay and satisfaction were those of Penner (1966), Lawler (1966), Reitz (1971), and Cherrington, Reitz, & Scott (1971). For both conceptual and methodological reasons, the evidence on the relation of satisfaction to performance-contingent pay remains somewhat equivocal even though persuasive arguments can be advanced for the expectation of a strong positive relationship. The first of these has to do with perceptions of the locus of causality on the part of the individual. (Rotter, 1966; DeCharms, 1968). That is, when pay and/or pay increases are contingent on performance, the individual has control over the amount that he receives to the extent that he has some control over his performance level. For any specific reward that this individual may want, therefore, the organization merely specifies the mechanism through which to obtain the reward. It is indeed the prerogative of the individual to decide upon the amount of the reward that he will choose to obtain. Of course, such control on the part of the employee exists not in a total sense, but within certain limits. For example, he can not increase performance infinitely. At the same time, however, the degree of control possible is greater than is the case with a system that pays its employees on the basis of, for example, seniority -- the employee can not control his age.

Various authors (e.g., de Charms, 1968) have argued that individuals seek to control their environment. Though the search for personal control is not the universal property of all individuals (e.g., Rotter, 1966; Lefcourt, 1966), manipulation of the environment to obtain desired ends should be pleasing to a greater or lesser degree. To the extent that it is within the power of the individual to obtain rewards from his job, he will be more satisfied. In contrast, when rewards are contingent on factors outside his control, he will be less satisfied with his job.

The second reason to expect performance-contingent rewards to be associated with higher satisfaction for employees is based on the notion of equity. To the extent that the financial rewards are provided contingent on good performance, they are less based on non-ability criteria such as, for example, sex, age, race. The outcomes for individuals are based to a large extent, therefore, on their inputs. In a social comparison situation, the input-outcome ratio for each individual would be approximately the same, resulting in feelings of equity (Adams & Rosenbaum, 1962; Adams, 1963). Such feelings of equity have been found to be related to the satisfaction of of employees (e.g., Pritchard, Dunnette, & Jorgenson, 1972). Garland (1973) found that need for money among subjects was not an important moderator, and that social comparison was in fact a significant variable in the determination and resolution of feelings of inequity. The findings of Lawler & O'Gara (1967) indicate that subjects adopted a number of techniques for dissonance reduction when they felt that they were being rewarded inequitably for their performance. Moore & Baron (1973) found manipulations of perceptions of qualifiedness (inputs) to have stronger effects on performance than manipulations of compensation (outcomes). These and other results obtained in the context of equity theory suggest social comparison to be a significant determinant of employees' affect toward their work and work situation. Such social comparisons are likely to result in more favorable estimates when rewards are based on 'fair' criteria. A potential problem with such comparisons may be that many people tend to exaggerate evaluation of their own performance (inputs) such that distorted perceptions of inequity may result (Meyer, 1975). At the same time, however, equitable social comparisons are more likely to be arrived at under a pay-for-performance system than under other systems. The equitable social comparisons would result in turn in greater feelings of satisfaction.

The third reason for expecting greater satisfaction under a performance-contingent reward system involves the concept of ambiguity. A reward system that bases outcomes on performance must necessarily employ a number of criteria of good performance. Such a system, therefore, tends to outline clear expectations as to role performance and performance assessment for organizational members, in contrast to one that rewards members on random explicit or variable implicit criteria. Role clarity has been found in a number of investigations to be positively related to job satisfaction (e.g., Beehr, 1974). To the extent, then, that a performance-based reward system provides more clarity in the work environment than one based on ambiguous criteria, employees will be more satisfied with such a system.

Finally, it may be argued that with respect to most extrinsic rewards, individuals seek to maximize the returns they wish to obtain. When rewards are provided uniformly or categorically on the basis, for example, of mere organizational membership, such individual maximization is not possible. Rewards that are contingent on performance, however, permit such maximization. The latter reward system would, therefore, be more satisfying.

Interaction between Performance-Contingent Pay and Intrinsic Rewards

Within the context of the cognitive evaluation theory, Deci (1971, 1972) has argued that when extrinsic rewards, specifically monetary rewards, are provided contingent on good performance of an intrinsically motivating task, the intrinsic motivation to perform the task is reduced. Despite the fact that Deci's results were concerned with the differential effects of intrinsic and extrinsic rewards on motivation rather than on satisfaction, it seems fruitful to examine the interaction between the two sets of rewards within the context of satisfaction.

Moderating Effects

In summary, then, this study proposed that there would be found a positive relationship between job satisfaction, on the one hand, and the degree to which the receiving of rewards is perceived to be contingent upon good job performance. Two classes of rewards will be distinguished: intrinsic rewards potentially inherent in performing the job activities and extrinsic rewards -- specifically amount of pay provided by others in connection with the job. It was further proposed that such a relationship may well be moderated by various situational and personal factors in such a way as to diminish or enhance the hypothesized relationship. The following paragraphs describe the seven potential moderators that were employed in the study.

Importance of rewards. Job satisfaction is a function, not only of what the individual gets from the job but also of what he wants from the job. In the present context, it appears likely that the allocation parameters are more important for those rewards that the individual values most. An individual who values pay highly but not job resources may care a great deal whether pay is based on seniority or performance, while being indifferent to the allocation criteria for resources. Facet-importance may, therefore, be a significant moderator of the contingent rewards-satisfaction relationship. It might be expected that the basic relationship here investigated will be stronger for short-tenure than for long-tenure employees.

Tolerance of ambiguity. A system that provides rewards contingent on performance may establish clear allocation parameters. Therefore, there should exist clarity regarding behaviors necessary for the receipt of valued rewards. Lyons (1971) showed that intolerance of ambiguity serves as a moderator of the relationship between such clarity in role prescription and job satisfaction. It is expected that persons with low tolerance for ambiguity will show a relatively stronger association between the performance contingency of rewards and job satisfaction.

Locus of control. Rotter (1966) postulated a 'control' construct that referred to a generalized expectancy about whether or not an individual has power over what happens to him/her. A system that allocates rewards on the basis of performance allows relative freedom to individuals to determine what they do or do not get. People who more strongly prefer to have control over their own life should be more satisfied under such a system than those who are resigned to, or who prefer, having external events control their lives.

Self-esteem. Self-esteem has been defined as the extent to which a person sees himself as a competent need-satisfying individual (Gelfand, 1962; Korman, 1970). To the extent that a performance-based reward system places an individual in a situation where he makes choices about his/her reward level, high self-esteem individuals may be more satisfied under such a system than will low self-esteem individuals.

Current level of performance. If rewards from organizational membership are provided contingent on performance, the performance level of the individual will determine the extent to which he/she actually receives these rewards. On the assumption that an individual would prefer to receive more of the rewards rather than less, it may be hypothesized that individuals will be more satisfied when they do receive the rewards.

Actual reward level. It has sometimes been argued that it is the actual reward that an individual receives not the contingency characteristics of such a reward, that is the relevant determinant of satisfaction. While there may be much truth to this argument, it is asserted here that the contingency characteristic of rewards will have an impact on satisfaction, above and beyond the effects of actual reward level.

Method

Data were obtained from 649 employees in five midwestern organizations including a hospital, a printing company, two companies that manufactured automotive accessories, and a research and development firm. The primary data source was a standardized interview conducted in the respondents' homes by professional interviewers. Additional data were obtained from on-the-job observations, employers' personnel records, and supervisors' ratings of subordinate performance. For a detailed discussion of the sample and data collection procedures, see Chapter 1.

Main Variables

The main variables employed in this analysis include measures of the degree to which rewards are contingent upon good job performance, and measures of job satisfaction. Two of the "contingency" measures were objective in derivation and others were derived from the reported perceptions of the respondents. The satisfaction measures were of several kinds as described later. A list of these variables appears as part of Table 1.

Perceived degree of reward contingency. Respondents indicated on three- or four-point scales the degree of certainty (likelihood) that specified rewards would occur if they did their jobs well. Example: "How likely is it that this may happen when you do your job well? . . . I will get a bonus or a pay increase". Single item measures were used for the rewards of pay, resources, freedom, and security in job. Multiple items were combined for the measures of intrinsic rewards, and of supervisory and co-worker approval.

Objective degree of reward contingency. Objective measures of contingency were feasible for two rewards: intrinsic job rewards and pay. The first was derived from the ratings of trained on-the-job observers with respect to four job characteristics: autonomy, variety, task identity, and feedback. These were combined in a multiplicative formula following a procedure suggested by Hackman & Lawler (1971). Further detail about the components of this index is provided in Chapter 4.

The second objective contingency measure was derived from knowledge of the pay practices in the firms, which allowed the ordered distinction between those employees in firms where: (1) Pay is fixed by contract or

Table 1

Descriptive Statistics on Variables Used in the Analyses

Variable	N	Range	Mean	Standard Deviation	Internal Consistency*
Perceived Degree of Performance Contingency					
Intrinsic Rewards	642	1 - 4	3.08	.98	(4) .87
Pay	644	1 - 3	2.47	.69	
Resources	644	1 - 3	2.26	.71	
Freedom	642	1 - 3	2.23	.70	
Security	642	1 - 3	1.89	.68	
Supervisor Approval	641	1 - 4	3.07	.78	(2) .64
Co-worker Approval	553	1 - 4	3.12	1.00	(2) .71
Objective Degree of Performance Contingency					
Intrinsic Rewards	578	1-10.9	2.58	2.12	
Pay	(three ordered but unscaled categories)				
Satisfaction					
Intrinsic Job Satisfaction	645	1 - 4	3.00	.77	(4) .82
Facet-free	646	1 - 5	3.94	.97	(4) .80 SB
Facet-specific	645	1 - 4	3.00	.59	(7) .86
Overall	645	1 - 4	2.92	.63	
Advancement	644	1 - 4	2.64	.92	
Pay	646	1 - 4	2.76	1.02	
Resource Adequacy	645	1 - 4	2.99	.64	(5) .79
Freedom	642	1 - 4	3.20	.78	
Security	644	1 - 4	3.13	.99	
Supervisory Approval	645	1 - 4	3.10	.82	(4) .88
Co-worker Approval	645	1 - 4	3.17	.58	(4) .84
Facet Importance					
Intrinsic Rewards	646	1 - 4	3.69	.41	(7) .78
Pay	646	1 - 4	3.56	.68	
Individual Differences					
Tenure	620	1 -40	8.36	7.03	
Self-esteem	645	1 - 7	6.01	1.05	
Intolerance of Ambiguity	642	1 - 5	2.83	1.27	(3) .45 SB
Locus of Control	635	1 - 2	1.63	.38	(2) .44 SB
Other					
Performance Rating	284	1 - 7	5.61	1.07	
Intrinsic Reward Level	648	1 - 5	3.68	.84	
Pay per Year (thousands)	647	3 -65	10.18	5.29	

*The internal consistency indicators are coefficient alphas except in cases where the Spearman-Brown (SB) coefficient was used. Numbers in parentheses indicate the number of component variables.

formal wage scale with few individual deviations; (2) Base pay is supplemented by a firm-wide bonus calibrated to base wage; and (3) Base pay is supplemented by group bonus based upon measured performance of the respondent's own work group. The first was judged to be least responsive to individual performance, and the third most responsive to individual performance. A number of respondents remained unclassified, as no objective grounds could be adduced for them.

Facet-specific job satisfactions. The interview included numerous questions asking for the respondent's degree of satisfaction (on four-point scales) with respect to various specific job facets. Using a hierarchical clustering method (Andersberg, 1973) satisfaction items or sets of items were created to parallel the "contingency" measures mentioned earlier. Measures resulted representing degree of satisfaction with: pay, intrinsic qualities of the job, supervisor, co-workers, resource adequacy, freedom, security, and advancement opportunities.

Non-specific job satisfaction. Three measures were derived to represent job satisfaction in a comprehensive sense. The first was derived by an unweighted average of the seven factor-specific satisfactions. The second was based upon the responses to four questions concerning the respondent's degree of job satisfaction phrased so that there was no reference to specific facets. Example: "Knowing what you know now, if you had to decide all over again whether to take the job you have now, what would you decide?" Further details on the derivation of this measure are provided by Beehr (1974) and in Chapter 3 of this report. A third measure, here called "overall job satisfaction", combines the facet-specific and facet-free measures.

Moderating Variables

Facet Importance. Twenty three items in the interview asked the respondent to rate on four-point scales ranging from "Not at all desirable" (1) to "Very desirable" (4) the desirability of various job facets on any job. Example: "How desirable is it . . . that you have a chance to develop your special abilities at work?" Seven of these items were chosen (on conceptual rather than empirical grounds) to represent the desirability of intrinsic rewards. The mean of these seven items was used as an index of the importance of intrinsic rewards. One item measured the desirability of pay to the respondents. For details on these two measures, see Gupta (1975).

Tenure was measured by the number of years the employee had been with the organization. It was categorized into three groups on the basis of the year of original hire: 1933-1964; 1965-1970; 1971-1973.

Tolerance of Ambiguity was an average of three items. Example: "What we are used to is always preferable to what is unfamiliar." The selection and and joint properties of these items are explicated in Beehr (1974).

Locus of Control. The internal-external locus of control index consisted of two items (Q No. 124, Volume II, Appendix A). Both items in this index referred to work-related perceptions of internal (self-controlling) and external (controlled by external events) locus of control. A high score on this index represented internal locus of control, and a low score represented external locus of control. The creation and properties of this index are described in Beehr (1974).

Self-Esteem was measured by three semantic-differential items concerning feelings of self-worth in the work role. This index was taken from Beehr (1974), who in turn used a modified version of the Quinn & Shepard (1974) self-esteem index.

Job Performance This was an index obtained from ratings by supervisors of employees along two dimensions, viz., the quality and the amount of work performed by his/her subordinates. These items were rated by supervisors on seven-point semantic differential scales (Volume II, Appendix E).

Intrinsic Reward Level. This concept was measured through the use of the challenge index from the overall quality of employment indicator (Barnowe, Mangione, & Quinn, 1972) and used elsewhere in this report (i.e., Chapter 3). It consisted essentially of the employee's descriptions of his job with respect to the extent to which characteristics that are relevant to intrinsic rewards were present. For a list of the constituent items, see Gupta (1975).

Actual Pay. Employees reported the actual amount of money they had earned from their job in the past year. This was used to measure their actual pay reward after conversion, as needed, to a full-time, full-year equivalent.

Results

The relationship between job satisfaction and the degree to which rewards are contingent on performance was examined in three ways.

First, one-way analyses of variance were performed using the extent to which each reward was contingent on performance as the predictor, and overall job satisfaction as the criterion. The results of these analyses are reported in Table 2. The table shows significant differences in the mean satisfaction levels for each of the reward areas. The largest differences occur in the areas of supervisory approval and intrinsic rewards, and the smallest in the area of coworker approval. Even the smallest differences are, however, significant beyond the .01 level. Not shown here are the

Table 2

Mean Satisfaction Levels for Groups Defined by the Perceived Performance Contingency of Different Types of Rewards

Type of Reward		Perceived Degree of Performance Contingency			
		Low	Medium	High	F
Intrinsic Reward	(N=639)	2.41	2.78	3.15	54.09*
Supervisory Approval [a]	(N=637)	2.25	2.65 3.03	3.20	48.28*
Co-worker Approval	(N=551)	2.68	2.87	3.04	10.62*
Resources	(N=641)	2.73	3.05	3.09	22.93*
Freedom	(N=636)	2.75	3.99	3.12	16.44*
Security	(N=639)	2.64	2.86	3.20	36.18*
Pay	(N=641)	2.79	3.06	3.21	22.28*

* $p < .01$.

[a] Four-point scale. The medium categories represent scale points 2 and 3.

results of analogous analysis of variance tests performed on all available combinations of rewards and alternative measures of satisfaction, as the results were without exception compatible with those shown in Table 2.

Also not shown here are analogous analyses performed using objective rather than perceived performance-reward contingencies as predictors, and satisfaction as the criterion. These analyses, however, provided a pattern of results similar to that shown in Table 2, although the effects were weaker. One reason for the relative weakness of the impact of objective contingency may be the questionable validity of the predictors.

The second test of the relationship between the contingency of rewards on performance and job satisfaction involved computing product moment correlations. The correlations were computed between (1) the extent to which different rewards were provided on the basis of performance, and (2) the relevant facet satisfaction measure, as well as facet-free satisfaction, total facet-specific satisfaction, and overall satisfaction. Table 3 reports these correlations. All correlations in this table were significant beyond the .01 level, indicating the existence of a positive relationship between the contingency of rewards on performance and job satisfaction.

The significance of the differences between the correlations of particular rewards with their corresponding facet-specific satisfactions and with the three more global satisfaction measures was computed using the formula for correlated samples (McNemar, 1969, p. 158). The tests revealed significantly higher correlations with facet-specific satisfactions than with the global satisfaction indices for three of the seven reward areas, *viz*., intrinsic rewards, supervisory approval, and resources.

Table 3

Correlations between Satisfaction and Perceived Degrees of Performance-Contingency of Rewards

	Perceived Contingencies						
	Intrinsic	Supervisory Approval	Co-worker Approval	Resources	Freedom	Security	Pay
Intrinsic Satisfaction	.41 (639)[a]						
Satisfaction with Supervisor		.67 (638)					
Satisfaction with Co-worker			.12 (551)				
Satisfaction with Resources				.12 (641)			
Satisfaction with Freedom					.23 (636)		
Satisfaction with Security						.27 (639)	
Satisfaction with Pay							.17 (632)
Facet-free Satisfaction	.29* (640)	.26* (639)	.17 (552)	.14 (642)	.17 (640)	.24 (641)	.19 (632)
Total Facet-specific Satisfaction	.29* (639)	.55* (637)	.17 (551)	.24° (641)	.19 (639)	.29 (640)	.27° (632)
Overall Satisfaction	.38 (639)	.45* (637)	.19 (551)	.24° (641)	.22 (639)	.32 (640)	.25° (632)

[a] Ns reported in parentheses. All correlations significant beyond the .01 level.

*Difference from correlation of relevant facet-specific satisfaction significant beyond .01 level.

°Difference from correlation of facet-specific satisfaction significant beyond the .05 level, direction reversed from prediction.

Finally, an index was made from the number of different rewards that were perceived to be very likely to follow from performance. A low score on this index indicated that only a few rewards were contingent on performance. One-way analyses of variance were then performed using this index as a predictor and two satisfaction variables -- facet free and overall -- as criteria. The results of these analyses are reported in Table 4. The table indicates that both types of satisfaction are higher when several rewards are contingent on performance than when only a few rewards are contingent.

The major hypothesis, that there exists a positive relationship between the contingency of rewards on performance and job satisfaction was, therefore, strongly confirmed.

Moderating effects. Several analyses were performed which looked at the moderating effects of individual differences, performance, and actual reward level on the main relationship, and at the interaction between intrinsic and pay rewards in the determination of satisfaction. The data from these analyses are not shown here, but the general results are summarized.

Individual differences. Individual differences provided weak and mixed results as moderators of the contingent rewards-satisfaction relationship. The two facet importance variables had some moderating effects of the expected kind, as did self-esteem. Locus of control had main, rather than interaction, effects. Hypotheses involving the moderating effects of tolerance of ambiguity and tenure were disconfirmed.

Performance. Hypotheses involving the moderating effects of current level of performance were disconfirmed. Performance was not a significant

Table 4

Mean Satisfaction for Groups Defined by Number of Rewards Perceived to be Contingent on Performance

Type of Satisfaction	Number of Contingent Rewards			
	Few	Medium	Many	F
Facet-free Satisfaction	3.71	3.93	4.19	11.35*
Overall Satisfaction	2.72	2.87	3.19	29.01*

*N = 538, p < .01

moderator of the relationship between performance-contingent rewards and satisfaction.

Actual reward level. It was hypothesized that a positive relationship would be detected between the contingency of intrinsic and pay rewards on performance and satisfaction even when the effects of actual reward level were held constant. This hypothesis was confirmed unequivocally with respect to intrinsic rewards; the data showed only slight support for the hypothesis with reference to pay.

Interaction between intrinsic and pay rewards. This hypothesis was not confirmed. The data showed that the provision of pay contingent on performance did not lower the intrinsic satisfaction of employees.

Discussion and Conclusions

The analyses confirmed that there existed in the data a direct relationship between performance-contingent rewards and satisfaction independent of the moderating effects of performance on actual reward level. In other words, the contingency of rewards is related to satisfaction not merely because of the receipt of rewards following good performance but because the perception of the contingency is in itself satisfying. Earlier, a variety of psychological processes that may cause such a relationship to exist were discussed. It was argued that equity, locus of control, and tolerance of ambiguity could explain the existence of a performance-contingent reward-satisfaction relationship. While the effects of perceptions of equity were not examined in this study, those of locus of control and tolerance of ambiguity were. The results showed that locus of control may be a psychological process involved in explaining the existence of such a relationship, but tolerance of ambiguity is not.

With reference to the two specific rewards that were given detailed treatment in the analyses, viz., pay and intrinsic rewards, an additive relationship emerged. Tests of Deci's (1971) hypothesis provided disconfirming evidence. One of the issues of relevance in instituting a performance-based monetary reward system is the psychological impact of such a system. If performance-based pay produces high levels of pay satisfaction but simultaneously has negative effects on other work-related attitudes, its overall utility is questionable. It is, therefore, encouraging that performance-contingent pay did not result in reduced intrinsic satisfaction. Three points should, however, be kept in mind when interpreting these results: (1) The present study was a cross-sectional field study, whereas Deci's results were obtained in a laboratory setting; (2) Deci used behavioral, rather than self-report measures of intrinsic motivation; and (3) The present study used intrinsic satisfaction as the criterion variable, whereas Deci used intrinsic motivation as his criterion.

A reward system that uses performance as the basis on which to reward employees affects the satisfaction of employees not only with particular facets of the job, but with the overall job as well. The results indicated that a similar pattern of relationships existed between performance-contingent rewards and facet-specific as well as facet-free and overall job satisfaction.

The present study has several limitations which may have attenuated the results. These limitations include the use of one-item measures, questionable validity of the objective pay contingency index and of the performance ratings, and the dichotomous form of some of the personality measures. In addition, there are problems regarding the dichotomizing of rewards into intrinsic and extrinsic areas, when they might better be conceptualized as a continuum. Finally, a major limitation of the study revolves around testing a causal model using cross-sectional data.

Implications

The results suggest that the institution of a performance-based reward system will have multiple benefits for the organization. In addition to the positive effects on the motivation, effort, and performance of employees reported in other research (summarized in Mitchell, 1974). Such a system had positive effects on job satisfaction. Job satisfaction has been found to be related to such employee behaviors as turnover, absenteeism, and lateness (e.g., Mangione, 1973), behaviors that are costly to the organization. The present results reinforce the belief that a performance-based reward system would be effective along several criteria of organizational success.

Furthermore, satisfaction of employees is a valued state in and of itself. Job dissatisfaction has been treated as an indicator of strain and poor mental health,(e.g., Beehr, 1974). Argyris (1964) has advocated the integration of the goals of the individual and the organization. A performance-based reward system would prove of considerable use in realizing this objective.

It is worth noting that in this population of typical persons, jobs and employers, the absolute level of performance-reward contingency is at best a modest one. It is rare that an individual sees an unequivocal connection between his/her performance and the various rewards that he/she obtains. The results of the present study suggest that if such a connection were clearly perceived by the employee, benefits would accrue along several criteria of organizational effectiveness.

References

Adams, J. S. Toward an understanding of equity. Journal of Abnormal and Social Psychology, 1963, 67, 422-436.

Adams, J. S., & Rosenbaum, W. B. The relationship of worker productivity to cognitive dissonance about wage inequities. Journal of Applied Psychology, 1962, 46, 161-164.

Andersberg, M. R. Cluster Analysis for Applications. New York: Academic Press, 1973.

Argyris, C. Integrating the Individual and the Organization. New York: Wiley, 1964.

Avery, R. D., & Mussio, S. J. A test of expectancy theory in a field setting using female clerical employees. Journal of Vocational Behavior, 1973 3, 421-432.

Barnowe, J. T., Mangione, T. W., & Quinn, R. P. The Relative Importance of Job Facets as Indicated by an Empirically Derived Model of Job Satisfaction. Ann Arbor: Survey Research Center, 1972.

Beehr, T. A. Role Ambiguity as a Role Stress: Some Moderating and Intervening Variables. Unpublished doctoral dissertation, The University of Michigan, 1974.

Cherrington, D. L., Reitz, H. J., & Scott, W. E., Jr. Effects of reward and contingent reinforcement on satisfaction and task performance. Journal of Applied Psychology, 1971, 55, 531-536.

Dachler, H. P., & Mobley, W. H. Construct validation of an instrumentality-expectancy-task-goal model of work motivation: Some theoretical boundary conditions. *Journal of Applied Psychology Monograph*, 1973, *58*, 397-418.

DeCharms, R. *Personal Causation*. New York: Academic Press, 1968.

Deci, E. L. Effects of laterally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, *18*, 105-115.

Deci, E. L. The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance*, 1972, *8*, 217-229.

Garland, H. The effects of piece-rate underpayment and overpayment on job performance: A test of equity theory with a new induction procedure. *Journal of Applied Social Psychology*, 1973, *3*, 325-334.

Gelfand, D. The influence of self-esteem on the rate of verbal conditioning and social matching behavior. *Journal of Abnormal and Social Psychology*, 1962, *65*, 259-265.

Georgopoulos, B. S., Mahoney, G. M., & Jones, M. W., Jr. A path-goal approach to productivity. *Journal of Applied Psychology*, 1957, *41*, 345-353.

Graen, G. B. Instrumentality theory of work motivation: Some experimental results and suggested modifications. *Journal of Applied Psychology Monograph*, 1969, *53*, No. 2, Part 2, 1-25.

Gupta, N. The Impact of Performance-Contingent Rewards on Job Satisfaction: Direct and Indirect Effects. Unpublished doctoral dissertation. The University of Michigan, 1975.

Hackman, J. R., & Lawler, E. E. III. Employee reactions to job characteristics. *Journal of Applied Psychology*, 1971, *55*, 259-286.

Herzberg, F., Mausner, B., & Snyderman, B. B. *The Motivation to Work*. New York: Wiley, 1959.

Jenkins, G. D., Jr., Nadler, D. A., Lawler, E. E. III, & Cammann, C. Standardized observations: An approach to measuring the nature of jobs. *Journal of Applied Psychology*, 1975, *60*, 171-181. See also Chapter 4.

Korman, A. K. Toward an hypothesis of work behavior. *Journal of Applied Psychology*, 1970, *54*, 31-41.

Lawler, E. E. III. Managers' attitudes toward how their pay is and should be determined. *Journal of Applied Psychology*, 1966, *50*, 273-279.

Lawler, E. E. III. *Pay and Organizational Effectiveness: A Psychological View*: New York: McGraw-Hill, 1971.

Lawler, E. E. III, & O'Gara, P. W. Effects of inequity produced by underpayment on work output, work quality, and attitudes toward the work. *Journal of Applied Psychology*, 1967, *51*, 403-410.

Lefcourt, H. M. Internal versus external control of reinforcement: A review. *Psychological Bulletin*, 1966, *65*,(4), 206-220.

Lyons, T. F. Role clarity, need for clarity, satisfaction, tension, and withdrawal. *Organizational Behavior and Human Performance*, 1971, *6*, 99-110.

Mangione, T. W. The Validity of Job Satisfaction. Unpublished doctoral dissertation, The University of Michigan, 1973

McNemar, Q. *Psychological Statistics*. New York: Wiley, 1969.

Meyer, H. The pay-for-performance dilemma. *Organizational Dynamics*, 1975, *3*, 39-50.

Mitchell, T. R. Expectancy modes of job satisfaction, occupational performance and effort: A theoretical, methodological, and empirical appraisal. *Psychological Bulletin*, 1974, *81*, 1053-1077.

Mitchell, T. R., & Albright, D. W. Expectancy theory predictions of the satisfaction, effort, performance, and retention of Naval aviation officers. Organizational Behavior and Human Performance, 1972, 8, 1-20.

Moore, L. M., & Baron, R. M. Effects of wage inequities on work attitudes and performance. Journal of Experimental Social Psychology, 1973, 9, 1-16.

Penner, D. D. A study of the causes and consequences of salary satisfaction. Crotonville, New York: General Electric Behavioral Research Service (mimeo), 1966.

Porter, L. W., & Lawler, E. E. III. Managerial Attitudes and Performance. Homewood: Irwin-Dorsey, 1968.

Pritchard, R. D., Dunnette, M. D., & Jorgenson, D. O. Effects of perceptions of equity and inequity on worker performance and satisfaction. Journal of Applied Psychology Monograph, 1972, 56, 75-94.

Quinn, R. P., & Shepard, L. J. The 1972-73 Quality of Employment Survey: Descriptive Statistics with Comparison Data from the 1969-70 Survey of Working Conditions. Ann Arbor: Survey Research Center, 1974.

Reitz, H. F. Managerial attitudes and perceived contingencies between performance and organizational response. Proceedings of the Academy of Management, August 1971, 227-238.

Rotter, J. B. Generalized expectancies for internal vs. external control of reinforcement. Psychological Monographs, 1966, 80, Whole No. 609, 1-28.

Skinner, B. F. Contingencies of Reinforcement: A Theoretical Analysis. New York: Appleton-Century-Croft, 1969.

Vroom, V. H. Work and Motivation. New York: Wiley, 1964.

Chapter 21

WORK AND NONWORK

by

Graham L. Staines

and

David Pagnucco

ABSTRACT

This analysis investigates the relationships between workers' experiences on and off the job. Of particular interest is the debate over two rival hypotheses: the "spillover" hypothesis and the "compensatory" hypothesis. The former argues that workers' experiences on the job carry over into the nonwork arena, and possibly vice versa, such that there develops a similarity in the patterning of work and nonwork life. The latter argues that the work situation is likely to be deficient in need fulfillment, at least in some respects, for most workers and that they will compensate for these deficiencies in their choices of leisure and family activities.

Data from the present study support the spillover hypothesis for the most part. Such support is reflected in the positive correlations between degree of involvement in work (measured subjectively) and degree of involvement in nonwork. Support is also shown in the positive correlations between general types of activities engaged in at work and a corresponding set of types of activities in nonwork. In addition, the relevant literature establishes a positive bond between subjective reactions to work and to leisure and family life. The one major exception to this pattern of spillover concerns physical effort on the job. Workers who expend a relatively great amount of physical effort at work are less involved in their nonwork activities.

Work and Nonwork

Part I--A Review of the Literature

Graham L. Staines

Introduction

Despite its magnitude, the literature connecting experiences on and off the job seems fragmented, uneven, and in need of conceptual integration. The literature includes an older tradition of research on the relationship between work and leisure (cf. Parker & Smith, 1976), a newer body of writings connecting work life and family life (cf. Aldous, Osmond & Hicks, 1977; Kanter, 1977; Pleck, 1976) and more limited discussions of the relationship between work and other domains of nonwork such as political participation (cf. Allardt, 1976). As noted, the literature linking work to each of these domains of nonwork has tended to be discrete, with few attempts at theoretical systematization. The search for more generalized patterns of association between work and nonwork seems warranted and timely, although there is no compelling reason why work life should bear the same relationship to each of the various domains of nonwork.

Role terminology helps to provide an integrative analysis of the relationship between work and nonwork. Work is a role, family life represents a cluster of roles (including spouse and parent), and leisure is also a role, although, of course, there are many important differences among these major life roles. Of relevance here are such role-related concepts as: degree of role involvement, types and level of role expectations, types of role activities, level of role competence, types of needs satisfied by a role, and subjective role reactions (cf. Nye, 1976).

It would in principle be possible to correlate measures of all role-related concepts concerning work with all measures of the same concepts concerning nonwork. The existing literature, however, does not cover many of the permutations so generated. The more modest but realistic task attempted here is to examine the literature linking work and nonwork roles on just three of the role-related concepts: degree of role involvement, types of role activities, and subjective role reactions. Degree of role involvement refers to objective factors such as time and energy invested in the job, range of activities undertaken, as well as subjective feelings of involvement or investment. Degree of involvement in nonwork activities invokes the same criteria as applied to leisure pursuits, family and home activities, organizational involvements, and so on. The types of activities that various jobs entail may be measured along such dimensions as: working with things versus ideas versus people (e.g., Kohn, 1969), degree of discretion or autonomy (e.g., Meissner, 1971; Torbert & Rogers, 1973), degree of complexity (Kohn, 1969; Kohn & Schooler, 1973), active versus passive, variable versus repetitive, etc. The same criteria may be used to classify types of nonwork activities. Subjective reactions to work experiences may be scored in a positive or negative direction along such common dimensions as satisfaction, enjoyment, number of work-related problems, and health-related reactions affecting physical or mental health. Again, the criteria applied to work may be applied equally effectively to experiences outside the work environment.

The categories in this tripartite conceptual scheme (degree of role involvement, types of role activities, and subjective role reactions) are not, of course, watertight. For example, degree of role involvement, measured subjectively, can come close to the notion of positive role reactions. When types of role activities concern such dimensions as active/passive, moreover, their distinctiveness from degree of involvement seems blurred. And when types of role activities tap such dimensions as autonomy, variety, and intrinsic interest, they shade into positive role reactions such as satisfaction and enjoyment, but for the most part these three conceptual categories work well.

The present review concentrates on two general and competing approaches to the relationship between work and nonwork that have appeared frequently in the literature.[1] The first general approach asserts a fundamental similarity between what occurs in the occupational environment and what transpires elsewhere. Generically termed the positive approach in this discussion, it appears in the literature under a variety of labels: spillover, extension, generalization, familiarity, identity, isomorphism. The second or negative approach proposes an inverse association between work and nonwork. It asserts that work experiences and nonwork experiences tend to be antithetical. The negative approach may be recognized in the literature by such terms as contrast, compensation, opposition, competition, and heteromorphism. As a formal

---

[1] e.g., Aldous, 1969, Bacon, 1975; Breer and Locke, 1965; Burch, 1969; Dumazedier, 1967; Grubb, 1975; Hagedorn & Labovitz, 1968; Havighurst, 1961; Kanter, 1977; Kelly, 1972; Kemper & Reichler, 1976; Kornhauser, 1965; London, Crandall & Seals, 1977; Meissner, 1971; Neulinger, 1974; Parker and Smith, 1976; Rapoport & Rapoport, 1965; Ridley, 1973; Salvo, 1969; Seeman, 1967; Shepard, 1974; Spreitzer & Snyder, 1974; Torbert & Rogers, 1973; Wilensky, 1960, 1961; Willmott, 1971.

alternative to these two major hypotheses, there is the null position according to which work and nonwork are unrelated, as is suggested by the terms _separateness_, _compartmentalization_, _segmentation_, and _neutrality_. The positive and negative approaches, along with their null alternative, have been offered in the literature as general statements of the relationship between work and nonwork. As stated, these wholly general formulations brook no distinctions among different nonwork roles, different role-related concepts, different facets of each concept, different subsamples of workers, and so on. As later sections of this paper make clear, however, these general perspectives sometimes need to be clarified, qualified, and revised. For example, they may need to be reformulated in conditional or contingent terms, the positive approach holding under one set of conditions, the negative approach under another set, and neither approach under yet a third set of circumstances (cf. Dubin, 1956; Parker & Smith, 1976). Such a contingency approach may not be as neat as the two overall approaches but neither does it run a similar risk of being too simple.

Possible future qualifications notwithstanding, the general positive and negative approaches provide a convenient framework for reviewing the evidence on the relationship between work and nonwork. As depicted in Table 1, the three role-related concepts as they concern the work role may be crossed with the same three concepts as they pertain to nonwork roles. Table 1 therefore contains nine separate (numbered) cells. Each cell contains correlations between dimensions of work and nonwork and several cells invite specific hypotheses about the relationship between work and nonwork, based on the positive and negative approaches. Cell one, for example, explores the relationship between degree of involvement in work and degree of

Table 1

Cells Containing Positive and Negative Hypotheses which link Work and Nonwork

Nonwork

Work	Degree of Involvement	Types of Activities	Subjective Reactions
Degree of Involvement	1 Mechanisms for Positive Hypothesis Integration Personality type Skills & abilities Cultural pressures Mechanisms for Negative Hypothesis Fixed sum of scarce resources Uniform & stable preferences	2 Mechanisms for Positive Hypothesis Integration Personality type (preferences) Skills & abilities Mechanism for Negative Hypothesis Desire for variety	3 No Mechanisms Cited for Positive Hypothesis Mechanisms for Negative Hypothesis Fixed sum of scarce resources Work as an alternative to traditional roles
Types of Activities	4 No Mechanisms Cited for Positive Hypothesis Mechanism for Negative Hypothesis Energy as scarce resource	5 Mechanisms for Positve Hypothesis Integration Skills & abilities Learned habits Strong preferences Mechanisms for Negative Hypothesis Energy as scarce resource Uniform & stable preferences Desire for variety	6 No Positive or Negative Hypotheses
Subjective Reactions	7 No Mechanisms Cited for Positive Hypothesis Mechanism for Negative Hypothesis Uniform & stable preferences	8 No Positive or Negative Hypotheses	9 Mechanisms for Positive Hypotheses Integration Skills & abilities Absence of work problems means family relationships unimpaired No Mechanisms Cited for Negative Hypothesis

involvement in nonwork roles. Cell two concerns associations between degree of work involvement and types of nonwork activities. Cell three considers connections between degree of work involvement and reactions to nonwork roles. Cell four correlates the types of work activities undertaken with degree of involvement in nonwork roles, and so on.

The present plan includes a set of procedures for several of the nine cells: presenting the specific hypotheses based on the positive and negative approaches; spelling out the causal mechanisms behind each hypothesis; and reviewing the available correlational evidence that bears on the hypotheses.

Consistent correlational patterns so generated raise the complex issue of causal direction. Do the correlations support the causal mechanisms according to which work affects nonwork, do they favor the mechanisms that explain why nonwork affects work, or do they suggest that both causal directions are operative simultaneously (cf. Breer & Locke, 1965; Dumazedier, 1967; Kanter, 1977; Kohn & Schooler, 1973)? Although every set of correlations raises the issue of causal direction anew, some general principles may be outlined. Breer and Locke (1965) and Meissner (1971) point out that the work environment lacks the flexibility or malleability of nonwork. This is particularly true of work and leisure. A worker rarely picks and chooses among jobs the way he may among leisure activities. He cannot modify the content of his work the way he can his recreation. Nor can he quit his job as readily as he can abandon a hobby. To the extent that the worker cannot select and shape his work environment, it becomes less plausible to attribute correlations between work and leisure to the impact of leisure on work. The causal connection between work and family appears more symmetrical than that between work and leisure. Family life may make demands on work which the work role must accommodate, just as work may require adjustments in family life (Kanter, 1977). Yet, despite the greater causal symmetry between work and family, the work experiences of employed people probably affect their family life more than the reverse.

Empirical Studies

Cell one

As applied to cell one, the positive approach predicts a positive correlation between degree of involvement in the work role and degree of involvement in nonwork roles. The causal arguments relevant here are fourfold. (1) Integration. Work and nonwork may be integrated, that is, interwoven or fused. Alternatively, a sharp demarcation may separate them (Aldous, 1969; Goldthorpe, Lockwood, Bechhofer, & Platt, 1969; Kelly, 1972; Parker & Smith, 1976; Ridley, 1973; Salaman, 1974). Integration means that a high degree of involvement in work accompanies a high degree of involvement in nonwork, and vice versa. As a causal mechanism for the positive approach, integration has at least four distinct components: overlap of time (work and nonwork activities undertaken simultaneously or in the same time period); overlap of place (work and nonwork activities occur in the same location); overlap of people (work and nonwork activities undertaken with the same people); and overlap of activities (certain activities have direct significance for both work and nonwork).[2] In occupational communities--the clearest example of integration of work and nonwork--work relationships, activities, interests and values permeate peoples' lives off the

---

[2] In addition to contributing to integration, a potential mechanism behind the positive hypothesis in cell one, overlap of activities provides direct evidence for the positive hypothesis in cell five (to be discussed shortly).

job. Members of such communities choose many of their friends from within their occupations, carry their work activities and interests into their leisure, belong to work-related clubs and societies, and define themselves largely in terms of their occupational role (Salaman, 1974). (2) Personality type. Consistent with the literature on the Type A personality (Rosenman, Friedman, Straus, Jenkins, Zyzanski, and Wurm, 1970; Rosenman, Friedman, Straus, Wurm, Jenkins, Messinger, Kositchek, Hahn, & Werthessen, 1966) is the notion that certain people have a general disposition to become heavily involved in all activities, including work-related as well as other ones. (3) Skills & abilities. Workers heavily involved in their jobs may acquire skills and abilities, especially social skills, that facilitate their involvement in nonwork activities (Hagedorn & Labovitz, 1968; Meissner, 1971). Alternatively, a similar but reversed causal arrow may run from skills and abilities acquired in nonwork roles to the work environment. (4) Cultural pressures. Certain segments of the working population (e.g., corporate executives) may experience pressures to become involved in social and political activities as well as work responsibilities (Levinson, Price, Munden, Mandl, & Solley, 1962).

The scholarly literature has exhibited limited interest in the positive hypothesis in cell one. Lipset, Trow, & Coleman (1956) reviewed some of the structural conditions of printers' jobs that, they believed, integrated work and nonwork activities in the printers' occupational community. They suggested that interest in the work sphere generalized to work-related but off-the-job political activity, namely, involvement in trade union activities. According to Allardt (1976), Seppänen (1958) obtained an empirical result along similar lines in a Finnish study: workers highly involved in their work (i.e., those desiring to improve occupational skills, unwilling to change factories, apt to make suggestions) proved particularly active in the unions. As regards off-the-job political activities without work ties, Lipsitz (1964) selected small samples of three occupational groups at a unionized assembly plant in New Jersey: semi-skilled assembly-line workers (N=12) who scored very low on degree of work involvement (91% would have left their present jobs if possible); semi-skilled repair, relief, and utility workers (N=14) who scored fairly low on involvement (71% would have left); and skilled maintenance workers (N=15) who exhibited high involvement (only 27% would have left). The three occupational groups, Lipsitz reported, scarcely differed at all on a variety of measures of extent of political participation (voting, discussion of politics, and interest in political affairs) but the small sample sizes limit the significance of his study.

The negative hypothesis in cell one, in contrast, has received very extensive attention in the literature. According to its most obvious formulation, the more time, energy, and money devoted to work activities, the less these resources are available for off-the-job activities (e.g., Clark, Nye and Gecas, 1977; Hammond, 1954). In short, such resources add to a fixed sum, and work and nonwork act as mutually exclusive alternatives vying for the allocation of scarce resources. While the argument about resources has general applicability for the negative hypothesis, its formulation often differs for the sexes. The argument for men stresses that heavy involvement in the work role may preclude or at least diminish participation in family and leisure activities. Discussions of women emphasize that major family commitments often conflict with job demands, thus forcing women either partially or completely out of the labor force or, failing that, forcing upon them hard personal decisions as to which set of role demands takes priority

(Bailyn, 1970; Cumming, Lazer and Chisholm, 1975; Safilios-Rothschild, 1971). Much of the literature on women also emphasizes that the incompatibility of work and nonwork roles applies particularly during those portions of the life-cycle when family roles make their maximum demands, that is, when young children (especially preschoolers) need constant attention at home (Harry, 1976; Meissner, Humphreys, Meis & Scheu, 1975).

Meissner (1971) suggested another possible mechanism behind the negative hypothesis in cell one. People, he claimed, have relatively uniform and stable preferences for levels (and types) of activity and involvement; their preferences have lower and upper bounds; and their preferences apply across major domains of activity and in the relatively short run. Thus, what people get from their experiences at work they do not need to seek outside of work, and, likewise, the reverse (Grubb, 1975; Hammond, 1954; Matthews & Abu-Laban, 1959).

Available empirical data have not adequately explored the negative hypothesis in cell one. The methodologies in relevant studies have frequently, though unnecessarily, presupposed the accuracy of the negative hypothesis (cf. examples of measures in Haavio-Mannila, 1971; Parker, 1965; Spreitzer & Snyder, 1974). The approach based on time budgets (e.g., Nelander, 1971), for example, has divided the hours of a day (or week) into a segment associated with work and segments associated with leisure activities, family roles, rest, etc. The time budget has thus built in a negative relationship between the amount of time allotted to work and the amount allotted to all nonwork activities, but not between time allotted to work and time allotted to particular nonwork roles. Studies by Walker and colleagues (Walker & Gauger, 1973; Walker and Woods, 1976) obtained negative associations among husbands between hours employed and hours devoted to household tasks and to parental role activity. Yet other recent studies by Clark and his colleagues (Clark & Gecas, 1977; Clark et al., 1977) found no significant relationship between husband's time in the work role and degree of participation in three family roles (housekeeping, [marital] therapeutic, and child socialization). Husbands' work time, however, did have a significantly negative effect on two other family roles, recreation with spouse, and childcare. Other studies have required forced choices between indicators of work and nonwork involvement, for example, Dubin's investigations of whether or not work is a central life-interest (e.g., Dubin, 1956). With few exceptions, available studies have failed to select criteria of degree of involvement (e.g., levels of energy, number of organizational affiliations, etc.) that have made the negative hypothesis in cell one a genuine and interesting issue. As to exceptions, Grubb (1975) related an eight-item measure of degree of work involvement (labeled affective boredom) to nine single-item measures of recreational participation (boating, camping, outdoor swimming, golf, etc.). Respondents had previously characterized these nine specific activities as more stimulating than their jobs. Based on a sample of hourly-paid production workers from four auto-assembly plants, Grubb found no clear evidence of the expected negative relationship between degree of work involvement and participation in a series of favorite leisure activities.

Cell two

As noted earlier, cell two concerns the relationship between degree of work involvement and types of nonwork activities. The positive approach suggests the hypothesis that more involved workers will choose leisure activities similar to or related to their work. It gains theoretical support from three of the positive mechanisms in cell one: integration, personality type (those inclined to become highly involved in work allow the work role to dictate their selection of nonwork activities), and skills and abilities acquired. Parker (1965) championed the positive hypothesis in his study of three occupations: bank employees, youth employment officers, and child care officers. Based on a measure of degree of intrinsic involvement in occupation (viz., choice of same/similar versus different occupation regardless of money), Parker ranked bank workers low (61% would have chosen a different occupation), youth employment personnel fairly high (20%) and child care officers very high (10%). Parker's data provided clear support for the positive hypothesis. High on degree of involvement, youth employment and child care officers frequently engaged in work-related forms of leisure; whereas bank officers, low on involvement, generally selected leisure activities completely different from and unconnected with their work. No discussions have appeared to date on the negative hypothesis that highly involved workers will pursue leisure activities that differ sharply from their occupational responsibilities, although a desire for variety in activities seems a plausible consideration here.

Cell three

Cell three explores the relationship between degree of work involvement and reactions to nonwork activities. According to the relevant positive hypothesis, extensive work involvement induces a favorable reaction to leisure and family, whereas the negative hypothesis posits that high work involvement creates dissatisfaction with off-the-job activities because of insufficient time and energy available (fixed sum of scarce resources).

The reverse causal direction whereby subjective reactions to leisure and family affect involvement in work is also a plausible possibility. Based on a sample of college-educated British couples, Rapoport, Rapoport & Thiessen (1974) found that husbands with a high degree of work involvement (i.e., those who reported more satisfaction from career than from family) enjoyed leisure activities less than did husbands who were family-oriented (no significance tests used).

In line with the notion of _scarce resources_, especially time and energy, the negative hypothesis gains plausibility in the case of working wives. Extensive involvement in the work role may arouse anxiety and guilt regarding their performance in traditional family roles. Alternatively, wives dissatisfied with their traditional roles may seek greater involvement in work (_work as an alternative to traditional roles_). As regards relevant empirical studies, Haavio-Mannila (1971) used a sample of employed wives from Helsinki, Finland, to test the relationship between degree of work involvement (seven noneconomic reasons for employment) and satisfaction with leisure, family and spouse. Whereas satisfaction with leisure bore no relationship to degree of involvement, satisfaction with family and with spouse both correlated negatively and, in several cases, significantly, with noneconomic motives for working. Ridley (1973) studied the relationship between job involvement (time devoted to occupational role performance in excess of the normal work day) and marital adjustment (a nine-item scale) among a sample from Florida of married female school teachers and their husbands. Despite several slight trends towards a negative association, Ridley obtained no significant relationship among either the teachers or their husbands.

Two studies have investigated the relationship between degree of involvement in work and reaction to the household role. Based on a sample of young, working Austrian women, Haller and Rosenmayr (1971) reported significantly negative associations between percent who wished to continue working and satisfaction with the household role. Similarly, Shea, Spitz, Zeller and Associates (1970) found that, for a national probability sample of women aged 30-44, employed white women who disliked housekeeping activities exhibited a stronger propensity to be in the labor force. The relationship did not hold, however, among the black women in their sample. Since, to a limited degree, women may view paid work as an alternative to housework, those who dislike housework will prefer greater involvement in the occupational role.

Cell four

Cell four introduces the correlations involving types of work activities and degree of nonwork involvement. Again appealing to the notion of scarce resources, particularly energy, the negative hypothesis in this cell links the extent to which the job makes heavy (and fatiguing) demands to a low degree of involvement in nonwork activities; the positive hypothesis proposes the reverse linkage. The available evidence on work and leisure slightly favors the negative hypothesis. Larrue (1965) detected no clearcut association between the level of fatigue induced by jobs among metal workers in Toulouse and the extent of their involvement in a variety of leisure activities, though two activities were more frequent among the fatigued workers: going to the movies, and going out into the country. Lundahl (1971) noted that employees in heavy and fatiguing jobs participated in fewer leisure activities, especially cultural and intellectual activities.

Cell five

Cell five concerns the widely discussed relationship between types of work and nonwork activities. The positive hypothesis predicts that workers choose nonwork activities that match in character what they do on their jobs. A number of possible causal mechanisms support the positive hypothesis. Integration of work and nonwork is clearly one relevant mechanism. Similarly, skills and abilities acquired during work activities may facilitate similar behavior off the job, and vice versa. Breer and Locke (1965) have emphasized that behavioral patterns learned and rewarded on the job (i.e., learned habits) generalize to nonwork activities, especially when based on many years of job experience (cf. Burch's [1969] discussion of "force of habit"). Strong preferences among workers for certain types of activities also underlie the positive hypothesis. These proposed mechanisms behind the positive hypothesis in cell five--integration, skills and abilities, habits, and strong preferences--may generate different results depending on whether nonwork activities primarily involve leisure or family life.

As regards work and leisure, Meissner (1971) tested the positive hypothesis on a sample of 206 male, Canadian, nonsupervisory workers employed in the forest industry and located in a small factory town on Vancouver Island. He focused on jobs that inhibited autonomy via technical constraints (viz., machine pacing, spatial confinement, task dependence, and type of work) and also on those that precluded the development of social skills through opportunities for social interaction. Though he used no tests of statistical significance, Meissner found, in line with the positive hypothesis, that jobs that imposed technical constraints or social isolation at work were negatively associated with leisure activities which required initiative, planning, decision-making and coordination (e.g., participation in voluntary organizations). He likewise found that jobs which precluded social interaction at work also related negatively to sociable activities off the job that took place outside an organizational context (e.g., visiting, talking, beer drinking, and going for family outings).

Also in line with the positive hypothesis, yet based on data at an aggregate as opposed to an individual level, Hagedorn and Labovitz (1968) noted that members of occupations that promoted leadership skills (percentage of all lower-status contacts required by the work role) scored higher on two types of social involvement off the job: belonging to, and participating in, community associations. Surprisingly, however, they also found that members of occupations that precluded social interaction on the job (percentage of work contacts that occurred infrequently) joined and participated in community associations more frequently.

The empirical studies relating work activities to family activities are more suggestive than definitive in their support of the positive hypothesis. Several investigations (e.g., Aberle & Naegele, 1952; Miller & Swanson, 1958) have reported that parents adopted approaches to child-rearing that were consistent with the values propagated in the father's work environment, such as competitiveness and aggressiveness. Yet these studies have contained little or no hard evidence of behavioral parallels between a father's activities at work and at home. A reversed version of the positive hypothesis has special significance for working women. The typically female occupations (e.g., school teacher, social worker, nurse, secretary) call for the same types of nurturant, supportive and interpersonal behavior that women regularly perform in their traditional family roles (Bailyn, 1964; Kanter, 1977). Traditional female socialization and role experiences may thus facilitate performance in these sex-typical occupations. In a more pessimistic vein, Agassi (1975, p. 292) has commented on the further parallel between paid employment and housework for many wives:

> "the domestic role demands very little intellectual exercise, so do most women's jobs; domestic work requires a lot of routine and repetition, women's jobs are the most repetitious in the labor market; the housewife's products are without permanence, women are concentrated in the production of nondurables; domestic work consists largely of personal cleaning and feeding services, women are concentrated in cleaning and feeding jobs; housewives suffer from lack of achievement and recognition, most women's jobs lead nowhere."

Few studies have empirically supported the negative hypothesis that workers choose leisure activities that differ from their work life, although three of the mechanisms cited earlier apply here: energy as a scarce resource, uniform and stable preferences, and a desire for variety. In a nonquantitative and impressionistic report of interviews with white- and blue-collar workers, for example, Gardell (1976) stressed that physically and mentally fatiguing jobs often permitted only passive pursuits (watching television, skimming newspapers) and domestic types of leisure (puttering about the home) which sometimes had to be preceded by a period of recovery from a strenuous work day.

Cell six

Cell six connects types of work behaviors and reactions to nonwork activities, but the general positive and negative approaches do not suggest any specific hypotheses pertaining to this cell.

Cell seven

Cell seven compares reactions to work with degree of nonwork involvement. Its positive hypothesis joins favorable work reactions to extensive nonwork involvement; its negative hypothesis does the opposite. Meissner (1971) has conveniently summarized three relevant nonAmerican studies (Larrue, 1965; Hanhart, 1964; Seeman, 1967), all of which supported the null rather than the positive hypothesis. Larrue (1965) who investigated the leisure of metal workers in Toulouse reported no significant relationship between work satisfaction and degree of involvement in a series of leisure activities. While not a quantitative report, Hanhart's (1964) study of workers in Zurich suggested that there was no association between work satisfaction and a series of leisure pursuits although dissatisfied workers did appear to go to the movies more frequently. Too recent to be included in Meissner's review, Grubb's (1975) investigation of auto workers produced no clearcut, linear relationship between a three-item measure of (cognitive) boredom and participation in recreational activities described by respondents as more stimulating than their jobs. Seeman (1967) found for a sample of male Swedish workers no connection between a scale of work alienation (also interpretable as an inverse measure of degree of job involvement) and extent of political knowledge.

One of the causal mechanisms cited earlier, uniform and stable preferences, supports the negative hypothesis. If workers satisfy their major needs at work, that is, they will not be motivated to seek additional gratifications via extensive involvement in nonwork roles. At least one study has supported the negative hypothesis. In the city of Kuwait (then a restrictive environment for foreign professionals), Matthews and Abu-Laban (1959) found that male primary school teachers from overseas who were satisfied with their jobs engaged in fewer different leisure pursuits and spent less time on leisure than those dissatisfied.

Cell eight

Cell eight correlates reactions to work with types of nonwork activities, although the entries in this cell cannot be conveniently subsumed under positive, negative, or null hypotheses. A common idea in the literature is that certain types of leisure are typically a response to a highly frustrating work environment. Wilensky (1960), for example, has proposed that dissatisfaction with work exhibits two major off-the-job correlates: individuation, a solitary life-style without significant interpersonal contacts on or off the job, and one that translates into apathy (eating and drinking alone, passive spectatorship) or explosive aggression (fighting, law violation); and family-home localism, a habitual withdrawal into the narrow circle of kinfolk and friends, which excludes ties with the larger society. Similarly, Bacon (1975) reviewed types of leisure that various authors have linked to a frustrated and alienated response to work: obsessive preoccupation with false consumer needs (Gorz, 1965); passive, home-centered,idle amusement (Baran & Sweezy, 1966); explosiveness or violence (e.g., intemperance, sexual licence, dangerous driving, aggression against minority groups).

Yet in his empirical survey of male workers in an English community, Bacon found no relationship between job alienation (or dissatisfaction) and any of the following: level of material prosperity (e.g., ownership of house, car, washing-machine, refrigerator, etc.) or material ambition (desire for same); degree of physically passive/home-centered leisure (e.g., listening to the radio, watching television, reading, resting) versus autonomous/creative leisure (hobbies, visiting pubs or churches, playing bingo, or motoring for pleasure). Though he did not present his results in quantitative form, Hanhart (1964) detected no relationship between work satisfaction and visiting, club membership, participation in active sports events, attendance at sports events, or listening to the radio; but work satisfaction did appear to relate negatively to moviegoing. In a quantitative study, Larrue (1965) likewise obtained no association between job satisfaction and involvement in five types of leisure activities (viz., reading, handywork, spectator sports, movies, and trips to the country). In their study of teachers in Kuwait, similarly, Matthews & Abu-Laban (1959) detected no significant difference in job satisfaction between those who participated in leisure activities with a group of others and those who participated alone. Nor did a difference in satisfaction emerge between those who selected leisure activities that emphasized communication and those who opted for other types of recreation.

Cell nine

Cell nine, the final cell, tests whether the relationship between reactions to work and nonwork roles is positive or negative. Consistent with the positive hypothesis regarding work and leisure and thus with the mechanisms of integration and skills and abilities, Kornhauser's (1965) survey of male Detroit factory workers obtained positive tetrachoric correlations between job satisfaction and satisfaction with leisure (0.26) and community (0.32). Haavio-Mannila's (1971) Finnish study included employed men in addition to the married employed women discussed earlier. Regrettably, her measure of satisfaction with work included housework as well as paid employment. Satisfaction with work, she found, correlated positively and significantly with satisfaction with leisure, for two groups (unmarried men and married women) but not for two others (married men and unmarried women). Campbell, Converse and Rodgers (1976) reported positive correlations from a 1971 national sample between work satisfaction and satisfaction with leisure (i.e., how spare time was spent) and friendship. They, too, treated housework as work for some of the women in their sample. London, Crandall and Seals (1977) noted that for a 1972 national sample of American adults correlations between items measuring job satisfaction and those measuring satisfaction with leisure ranged from 0.01 to 0.28, with a median of 0.14. The strongest correlations were between satisfaction with coworkers and satisfaction with social life off the job; the weakest linked satisfaction with financial aspects and physical working conditions on the job to satisfaction with organizations belonged to off the job. London et al. argued, however, that shared methods variance may have accounted for these positive correlations.

Kemper and Reichler (1976) spelled out several plausible reasons why favorable reactions to marital and family roles among men might accompany high levels of job satisfaction. They noted that the husband satisfied with his work does not bring his troubles home to burden his wife and family; his positive frame of mind makes him more receptive to their needs; his achievements at work add to their admiration for him. Based on college students' reports of their parents, Kemper and Reichler's own study found significantly positive correlations between the father's intrinsic job satisfaction (work was meaningful/important, enjoyed work/liked to talk about it) and his marital satisfaction, but no relationship between his satisfaction with extrinsic aspects of the job (income, recognition) and his marital satisfaction. Separate studies by Bradburn and Caplovitz (1965), Renne (1970), Haavio-Mannila (1971), and Ridley (1973) all found that husbands dissatisfied with their jobs registered lower levels of marital adjustment. And satisfaction with job and with family life proved positively associated in several studies: Kornhauser (1956), husbands only; Haavio-Mannila (1971), significant correlations for both married and unmarried men; and Campbell et al. (1976), sample included both sexes.

Yet data on reactions to work and family roles among women suggest a more complicated pattern. Renne (1970) reported that satisfaction with job and with marriage were positively related among wives, though not as consistently as among husbands (no significance tests used). In Ridley's (1973) study of married female school teachers, he obtained zero association between job satisfaction and marital adjustment. For three subgroups (wives with school aged children; those whose husbands had education that was equal to or less than theirs; and those whose husbands belonged to a semiprofessional occupation), however, the association was significantly positive. Haavio-Mannila's (1971) study produced positive and significant correlations between work satisfaction and satisfaction with family among married but not unmarried women; yet no significant correlation emerged between satisfaction with work and with marriage (relationship to spouse) among married women.

Summary and Conclusions

As noted, the specific hypotheses reviewed here were drawn from two general approaches to the relationship between work and nonwork roles: a positive approach that stressed the similarity of work and nonwork, and a negative one that emphasized their dissimilarity. The two general approaches suggested specific positive and negative hypotheses for some of the nine cells in Table 1. For cells one, five, and nine, the general notions of similarity and dissimilarity logically implied positive and negative correlations, respectively, between work and nonwork roles as regards three role-related concepts: degree of role involvement, types of role activities, and subjective role reactions.

Several other cells contained equally interesting hypotheses which were also labeled positive and negative even though they could not be derived from the notions of similarity and dissimilarity without additional assumptions or qualifications. For example, in cells three and seven, which concerned degree of involvement in the work role and subjective reactions to nonwork roles (and vice versa), the derivations of the specific hypotheses required the further assumption that a high degree of role involvement and favorable role reactions both reflected a positive role orientation. As regards cell four (types of work activities and degree of nonwork involvement), one dimension of work activities (viz., whether work was heavy and tiring) was treated as a proxy for degree of work involvement so that the argument about scarce resources could be invoked as in cell one (i.e., heavy work discourages a high degree of nonwork involvement). Cell two required a more tenuous derivation. The positive prediction that a high degree of work involvement elicits work-related off-job behaviors assumed that the similarity between work and nonwork was conditional upon a high degree of work involvement rather than being universal. Virtually the same hypothesis could have been derived by an alternative, though equally tenuous, route involving cell five. Specifically, the positive prediction in cell five (viz., that similar behaviors are exhibited on and off the job) could have been arbitrarily restricted to workers with a high degree of work involvement. Cells six and eight, finally, contained no specific predictions derivable either directly or indirectly from the notions of similarity and dissimilarity.

The main inference to be drawn here is that the notions of similarity and dissimilarity as applied to work and nonwork roles generated a number of specific positive and negative hypotheses, some derived directly, others just as interesting but requiring additional assumptions or qualifications. Since the specific hypotheses in the various cells differed considerably, the evidence relevant to each hypothesis should be evaluated separately.

Thus, the many foregoing empirical tests of the positive and negative hypotheses call for a careful summary of the literature on a cell by cell basis. From a methodological standpoint, a large number of the studies contained major weaknesses: few took into account the resources needed for nonwork activities; few imposed other secondary controls on the relationship between work and nonwork; few explored such measurement problems as shared method variance; and several of the studies used measures whose categorization as degree of role involvement, types of role activities, or subjective role reactions, raised ambiguities. Perhaps partly because of methodological problems, the overall empirical pattern was not consistent and therefore conformed more closely to the segmentation argument (null hypothesis) than to either the generalization (positive hypothesis) or compensation argument (negative hypothesis). The summary of empirical findings that follows gives greatest weight to the studies with less problematic designs:

(1) Cell one initiated the pattern of inconsistent and nondefinitive findings. Grubb (1975) found no relationship between measures of the degree of involvement in work and leisure. A number of studies of husbands (Clark and Gecas, 1977; Clark et al., 1977; Walker and Gauger, 1973; Walker and Woods, 1976) reported a negative relationship between time spent working and time spent in certain (but not all) family roles. In addition, Seppänen (1958) obtained a positive association between degree of involvement in work and in work-related but off-the-job political activity.

(2) The second cell contained only one empirical study (Parker, 1965), and it supported the positive hypothesis that more involved workers choose leisure activities similar to their tasks at work.

(3) Cell three's studies of the relationship between degree of work involvement and subjective reactions to nonwork produced somewhat confusing results. Rapoport et al. (1974) obtained a negative association between degree of work involvement and enjoyment of leisure among a sample of educated British husbands. Haavio-Mannila (1971) obtained zero association among married women between job involvement and satisfaction with leisure but negative correlations with satisfaction with family and with spouse. Yet Ridley (1973) uncovered no association between degree of job involvement and marital adjustment for either men or women. Two

studies of women (Haller & Rosenmayr, 1971; Shea et al., 1970) recorded some evidence of a negative association between involvement in work and attitude to the household role.

(4) Cell four included mildly conflicting findings from two tests of the association between types of work activities and degree of involvement in leisure. Whereas Larrue (1965) obtained a marginally positive bond between fatiguing jobs and more leisure activities, Lundahl (1971) detected a reverse trend.

(5) Cell five's concern with the parallel between types of activities on and off the job produced another set of mixed results. According to Meissner's (1971) investigation, jobs that offered workers little autonomy were negatively associated with leisure that required planning and coordination; and jobs requiring interpersonal interaction correlated positively with leisure activities with strong social components. Yet Hagedorn and Labovitz (1968) recorded one diametrically opposite finding as well as a somewhat similar one.

(6) Cell six did not concern the positive or negative hypotheses.

(7) Cell seven compared reactions to work and extent of nonwork involvement, and included both zero and negative findings. Larrue (1965) and Grubb (1975) both reported no association between job satisfaction and leisure involvement. One other study (Matthews and Abu-Laban, 1959), however, uncovered a negative association between job satisfaction and leisure involvement. Seeman (1967) found no connection between work alienation and extent of political knowledge.

(8) Cell eight did not concern the positive or negative hypotheses.

(9) Cell nine contained two studies (Kornhauser, 1965; London et al., 1977) which obtained positive correlations between job satisfaction and satisfaction with leisure. London et al. (1977) have pointed out, however, that shared method variance may have accounted for low positive correlations between similarly phrased measures of satisfaction with work and leisure.

Though the qualifying comment about shared method variance applies again, cell nine also contained a number of entries reporting positive correlations between job satisfaction and satisfaction with family roles among husbands: Bradburn and Caplovitz (1965), job satisfaction and marital adjustment; Kemper and Reichler (1976), intrinsic job satisfaction and marital satisfaction; Kornhauser (1965), job satisfaction and satisfaction with family and home life; Renne (1970), job satisfaction and marital adjustment. In comparison, Ridley's (1973) study of wives obtained no association between job satisfaction and marital adjustment.

The foregoing review of the general positive and negative approaches to the relationship between work and nonwork should not obscure a variety of alternative research strategies that hold promise. A number of studies, for instance, have pursued more specific links between dimensions of work and nonwork. In her summary of the literature, Kanter (1977) cited a number of specific, occupationally-based patterns or events that have major implications for family life (e.g., shift work, on-the-job travel, and geographical relocation) as well as a number of specific ways in which family life can affect occupational experiences. A large number of investigations have explored the effects of work on nonwork by comparing individuals who work with those who do not. The employment status of wives has been linked to their marital adjustment (e.g., Staines, Pleck, Shepard & O'Connor, in press), their influence in family decisions (Blood & Hamblin, 1958), and their level of participation in political life (Andersen, 1975). Pursuing another line of inquiry, Champoux (1977; in press) has treated similarity between work and nonwork as an individual difference among employees rather than as an aggregate characteristic of the population of workers, and has used bipolar rating scales to allow workers themselves to define what is similar about work and nonwork and what is different.

References

Aberle, D. F. and Naegele, K. Middle-class fathers' occupational role and attitudes toward children. American Journal of Orthopsychiatry, 1952, 22, 366-378.

Agassi, J. B. The quality of women's working life. In: L. E. Davis and A. B. Cherns, (Eds.), The quality of working life, Vol. 1 New York: Free Press, 1975.

Aldous, J. Occupational characteristics and males' role performance in the family. Journal of Marriage and the Family, 1969, 31, 707-712.

Aldous, J., Osmond, M., and Hicks, M. Men's work and men's families. In: W. Burr, R. Hill, I. Nye, and I. Reiss (Eds.), Contemporary theories about the family. New York: Free Press, 1977.

Allardt, E. Work and political behavior. In: R. Dubin (Ed.), Handbook of Work, Organization, and Society. Chicago: Rand McNally, 1976.

Andersen, K. Working women and political participation, 1952-1972. American Journal of Political Science, 1975, 19, 439-453.

Bacon, A. W. Leisure and the alienated worker: A critical reassessment of three radical theories of work and leisure. Journal of Leisure Research, 1975, 7, 179-190.

Bailyn, L. Notes on the role of choice in the psychology of professional women. Daedalus, 1964, 93, 700-710.

Bailyn, L. Career and family orientations of husbands and wives in relation to marital happiness. Human Relations, 1970, 23, 97-113.

Baran, P. A. and Sweezy, P. M. Monopoly Capital. New York: Pelican, 1966.

Blood, R. O., Jr., and Hamblin, R. L. The effect of the wife's employment on the family power structure. Social Forces, 1958, 36, 347-352.

Bradburn, N. M. and Caplovitz, D. Reports on happiness. Chicago: Aldine, 1965.

Breer, P. E. and Locke, E. A. Task experience as a source of attitudes. Homewood, Ill.: Dorsey, 1965.

Burch, W. R. The social circles of leisure: Competing explanations. Journal of Leisure Research, 1969, 1, 125-147.

Campbell, A., Converse, P. E., and Rodgers, W. L. The quality of American life: Perceptions, evaluations, and satisfactions. New York: Russell Sage, 1976.

Champoux, J. E. Perceptions of work and nonwork: A re-examination of the compensatory and spillover model. Unpublished paper, University of New Mexico, 1977.

Champoux, J. E. Work, central life interests, and self-concept. Pacific Sociological Review, in press.

Clark, R. A. and Gecas, V. The employed father in America: A role competition analysis. Paper presented to the 1977 Annual Meeting of the Pacific Sociological Association, 1977.

Clark, R. A., Nye, F. I., and Gecas, V. "Career widows"? Husbands' work involvement and marital role performance. Paper presented to the 1977 Annual Meeting of the American Sociological Association, 1977.

Cumming, E., Lazer, C., and Chisholm, L. Suicide as an index of role strain among employed and not employed married women in British Columbia. Canadian Review of Sociology and Anthropology, 1975, 12, 462-470.

Dubin, R. Industrial workers' worlds: A study of the 'central life interests' of industrial workers. Social Problems, 1956, 3, 131-142.

Dumazedier, J. Toward a society of leisure. New York: Collier-Macmillan, 1967.

Gardell, B. Reactions at work and their influence on nonwork activities: An analysis of a sociopolitical problem in affluent societies. Human Relations, 1976, 29, 885-904.

Goldthorpe, J. H., Lockwood, D., Bechhofer, F., and Platt, J. The affluent worker in the class structure. London: Cambridge University Press, 1969.

Gorz, A. Work and consumption. In: P. Anderson and R. Blackburn (Eds.), Towards socialism. London: Fontana Library, 1965.

Grubb, E. A. Assembly line boredom and individual differences in recreation participation. Journal of Leisure Research, 1975, 7, 256-269.

Haavio-Mannila, E. Satisfaction with family, work, leisure and life among men and women. Human Relations, 1971, 24, 585-601.

Hagedorn, R. and Labovitz, S. Participation in community associations by occupation: A test of three theories. American Sociological Review, 1968, 33, 272-283.

Haller, M. and Rosenmayr, L. The pluridimensionality of work commitment. Human Relations, 1971, 24, 501-518.

Hammond, S. B. Class and family, Part I. In: O. A. Oeser and S. B. Hammond (Eds.), Social structure and personality in a city. London: Routledge and Kegan Paul, 1954.

Hanhart D. Arbeiter in der Freizeit: Eine sozial-psychologische Untersuchung. Bern: Hans Huber, 1964.

Harry, J. Evolving sources of happiness for men over the life cycle: A structural analysis. Journal of Marriage and the Family, 1976, 38, 289-296.

Havighurst, R. J. The nature and values of meaningful free-time activity. In: R. W. Kleemeier (Ed.), Aging and leisure. New York: Oxford University Press, 1961.

Kanter, R. M. Work and family in the United States: A critical review and agenda for research and policy. New York: Russell Sage Foundation, 1977.

Kelly, J. R. Work and leisure: A simplified paradigm. Journal of Leisure Research, 1972, 4, 50-63.

Kemper, T. D. and Reichler, M. L. Work integration, marital satisfaction, and conjugal power. Human Relations, 1976, 29, 929-944.

Kohn, M. L. Class and conformity: A study in values. Homewood, Ill.: Dorsey, 1969.

Kohn, M. L. and Schooler, C. Occupational experience and psychological functioning: An assessment of reciprocal effects. American Sociological Review, 1973, 38, 97-118.

Kornhauser, A. Mental health of the industrial worker. New York: Wiley, 1965.

Larrue, J. Loisirs ouvriers chez les Metallurgistes Toulousains. Paris: Mouton, 1965.

Levinson, H., Price, C. R., Munden, K. J., Mandl, H. J., and Solley, C. M. Men, management, and mental health. Cambridge, Mass.: Harvard University Press, 1962.

Lipset, S. M., Trow, M. A., and Coleman, J. S. Union democracy: The internal politics of the international typographical union. Glencoe: The Free Press, 1956.

Lipsitz, L. Work life and political attitudes: A study of manual workers. American Political Science Review, 1964, 58, 951-962.

London, M., Crandall, R., and Seals, G. W. The contribution of job and leisure satisfaction to quality of life. Journal of Applied Psychology, 1977, 62, 328-334.

Lundahl, A. Leisure and recreation: Low income investigation. Stockholm: Allmanna Forlaget, 1971.

Matthews, A. T. V. and Abu-Laban, B. Job satisfaction and leisure-time activity in a restrictive environment. Sociology and Social Research, 1959, 43, 189-196.

Meissner, M. The long arm of the job: A study of work and leisure. Industrial Relations, 1971, 10, 239-260.

Meissner, M., Humphreys, E. W., Meis, S. M., and Scheu, W. J. No exit for wives: Sexual division of labour and the cumulation of household demands. *Canadian Review of Sociology and Anthropology*, 1975, *12*, 424-439.

Miller, D. R. and Swanson, G. E. *The changing American parent*. New York: Wiley, 1958.

Nelander, S. *"Real" working hours of Swedish wage-earners: Low income investigation*. Stockholm: Allmanna Forlaget, 1971.

Neulinger, J. *The psychology of leisure*. Springfield, Ill.: Thomas, 1974.

Nye, F. I. *Role structure and analysis of the family*. Beverly Hills, Calif.: Sage Publications, 1976.

Parker, S. R. Work and non-work in three occupations. *Sociological Review*, 1965, *13*, 65-75.

Parker, S. R., and Smith, M. A. Work and leisure. In: R. Dubin (Ed.), *Handbook of work, organization, and society*. Chicago: Rand McNally, 1976.

Pleck, J. H. Men's new roles in the family: Housework and childcare. Unpublished paper, 1976. To appear in C. Safilios-Rothschild (Ed.), *Family and Sex Roles*.

Rapoport, R. and Rapoport, R. Work and family in contemporary society. *American Sociological Review*, 1965, *30*, 381-394.

Rapoport, R., Rapoport, R., and Thiessen, V. Couple symmetry and enjoyment. *Journal of Marriage and Family*, 1974, *36*, 588-591.

Renne, K. S. Correlates of dissatisfaction in marriage. *Journal of Marriage and the Family*, 1970, *32*, 54-67.

Ridley, C. A. Exploring the impact of work satisfaction and involvement on marital interaction when both partners are employed. *Journal of Marriage and the Family*, 1973, *35*, 229-237.

Rosenman, R. H., Friedman, M., Straus, R., Jenkins, C. D., Zyzanski, S., and Wurm, M. Coronary heart disease in the Western Collaborative Group Study: A follow-up experience of 4 1/2 years. *Journal of Chronic Diseases*, 1970, *23*, 173-190.

Rosenman, R. H., Friedman, M., Straus, R., Wurm, M., Jenkins, C. D., Messinger, H. B., Kositchek, R., Hahn, W., and Werthessen, N. T. Coronary heart disease in the Western Collaborative Group Study. *Journal of the American Medical Association*, 1966, *195*, 86-92.

Safilios-Rothschild, C. Towards the conceptualization and measurement of work commitment. Human Relations, 1971, 24, 489-493.

Salaman, G. Community and occupation. London: Cambridge University Press, 1974.

Salvo, V. J. Familial and occupational roles in a technological society. In: Perrucci, R., and Gertl, J. (Eds.), The engineers and the social system. New York: Wiley, 1969.

Seeman, M. On the personal consequences of alienation in work. American Sociological Review, 1967, 32, 273-285.

Seppänen, P. Dual allegiance to company and union. Helsinki: WSOY, 1958.

Shea, J. R., Spitz, R. S., Zeller, F. A., and Associates. Dual Careers: A longitudinal study of labor market experience of women, Vol. 1, Manpower Research Monograph no. 21 (Washington: U. S. Government Printing Office, 1970).

Shepard, J. M. A status recognition model of work-leisure relationships. Journal of Leisure Research, 1974, 6, 58-63.

Spreitzer, E. A. and Snyder, E. E. Work orientation, meaning of leisure and mental health. Journal of Leisure Research, 1974, 6, 207-219.

Staines, G. L., Pleck, J. H., Shepard, L., and O'Connor, P. Wives' employment and marital adjustment: Yet another look. Psychology of Women Quarterly, in press.

Torbert, W. R. and Rogers, M. P. Being for the most part puppets: The interaction of men's labor, leisure, and politics. Cambridge, Mass.: Schenkman. 1973.

Walker, K., and Gauger, W. Time and its dollar value in household work. Family Economics Review, Fall, 1973 8-13.

Walker, K., and Woods, M. Time use: A measure of household production of family goods and services. Washington, D.C.: American Home Economics Association.

Wilensky, H. L. Work, careers, and social integration. International Social Science Journal, 1960, 12, 543-560.

Wilensky, H. L. Life cycle, work situation, and participation in formal associations. In: R. W. Kleemeier (Ed.), Aging and leisure. New York: Oxford University Press, 1961.

Willmott, P. Family, work and leisure conflicts among male employees: Some preliminary findings. Human Relations, 1971, 24, 575-584.

Work and Nonwork:

Part II: An Empirical Study

Graham L. Staines and David Pagnucco

The present empirical study investigates six of the nine cells cited in Table 1 (Part I). It includes measures of degree of involvement in, types of activities in, and subjective reactions to, the work role, as well as measures of degree of involvement in and types of activities in the nonwork sphere. Lacking a measure of subjective reactions to nonwork roles, it addresses only cells one, two, four, five, seven and eight. Wherever appropriate, the analyses in this study pit the positive against the negative hypothesis. They also examine certain other relationships between work and nonwork that lie outside the directional hypotheses.

Sample

The sample for the present report consists of 651 employees from five organizations: a hospital, a printing company, a research and development laboratory, and two plants that manufactured automobile accessories--all in the midwest. Though not representative of any theoretical population, the respondents do represent a variety of common occupations within common kinds of work establishments. The sample includes both supervisory and nonsupervisory employees, but excludes those working fewer than 20 hours per week and those too new to their jobs to answer job-related questions. Compared to the national population of workers, the sample contains a higher proportion of blacks, women, young people, and those who have never married. A fuller description of the sample appears in Chapter 1.

The data reported in the present paper come from personal, standardized interviews with respondents during the winter months of 1972-73 (i.e., Phase I of the total study). The overall interview response rate for the five organizations was 72.9%. The personal interviews included questions concerning the worker's job values, job attributes, job attitudes, physical and mental health, social and political activities, personality, and personal background.

Measures

Degree of Work Involvement

The interview schedule included 12 single-item measures of degree of work involvement, three of them relatively objective and the remainng nine more subjective. In the list of these items that follow, responses indicating the greatest degree of involvement appear in italics.

Objective Measures

Hours worked per week: During the average work week how many hours do you work not counting time off for meals? (High number of hours indicates high involvement.)

Have second job: Do you presently have any other job or do any other work for pay outside of (study employer)? (*Yes*, No.)

Hours worked on second job: About how many hours a week on the average do you work for pay outside of (study employer)? (High number of hours indicates high involvement.)

Subjective Measures

Work harder than coworkers: Would you say you *work harder,* less hard or about the same as other people doing your type of work?

Frequency of extra work: How often do you do some extra work for your job which isn't required of you? (*Often,* sometimes, rarely, never.)

Do my best at work: How do you see yourself in your work? (*Doing my best,* . . . not doing my best.)

Work my hardest at work: How do you see yourself in your work? (*Working my hardest,* . . . not working hard.)

Would work even if enough money: If you had enough money to live as comfortably as you'd like, would you continue to work? (*Yes,* no.)

Time drags at work: On most days on your job, how often does time seem to drag for you? (Often, sometimes, rarely, *never.)*

Ideal job if free to choose: If you were free to go into any type of job you wanted, what would your choice be? (*Same job,* some other job, retirement, or simply not work.)

Job skills useful in future: How useful and valuable will your present job skills be five years from now? (*Very useful and valuable,* somewhat, a little, not at all.)

Degree absorbed in job: Some people are completely involved in their job--they are absorbed in it night and day. For other people, their job is simply one of several interests. How involved do you feel in your job? (Very little, slightly, moderately, *strongly involved.)*

Types of Work Activities

The indices of Quality of Employment provide the major measures of types of work activities. These include the overall index of Quality of Employment (31 component items), as well as more specific measures of quality of employment for four areas: resources (five items), financial rewards (five items), challenge (12 items), and comfort (nine items). Full details of these five measures of Quality of Employment appear in Chapter 3.

Several other items and indices cover somewhat different areas:

Always a lot of work to do: There is always a great deal of work to be done (Not at all true, a little true, somewhat true, very true.)

Requires a lot of mental effort: My job requires a lot of mental effort (as above).

Requires learning new things: My job requires that I keep learning new things (as above).

Requires creativity: My job requires that I be creative (as above).

Requires a lot of physical effort: My job requires me to exert a lot of physical effort (as above).

One person job: My job is pretty much of a one person job--there is little need for checking or meeting with others (as above).

Works in a group: Is there any group of people at (study employer) that you think of as your co-workers--people whom you see just about every day and with whom you have to work closely in order to do your job well? (Yes, no.)

Number in workgroup: About how many people are there in this group?

Workgroup cohesion: A mean score on a 10-item index concerning the positive qualities of the respondent's coworkers; whether the coworkers: help each other, use new ideas, get along together, react to unusual demands, finish what they set out to do, stick together, share news, do their own jobs, give respondent help, and defend each other from outside criticism.

Work pace determined by others: A mean score on a four-item index measuring whether the respondent's work pace is determined by other people (viz., workgroup, immediate superior, other groups or departments, customers or clients/patients).

Does respondent supervise: Do you supervise anybody as part of your job? (Yes, no.)

Number of workers respondent supervises: About how many people do you supervise directly?

Measures of Subjective Reactions to Work

Of the nine indices of subjective reactions to the work role, seven measure aspects of job satisfaction. Scores on all indices represent means of the component items; and, where appropriate, the response indicating the most positive reaction appears in italics.

Facet-free job satisfaction

Five single items tap overall job satisfaction: a rating of general job satisfaction; a question about whether the job measures up to the worker's original expectations; a question about whether the respondent would select the same job if he or she had to decide all over again; an item about whether the worker would recommend the job to a friend; and a question about whether the respondent would select something different as an ideal job.

Satisfaction with coworkers

The respondent indicates that he or she is not at all, a little, somewhat, or *very satisfied* with two components of the job: the friendliness of the people he or she works with; and the chance the job provides him or her to make friends.

Satisfaction with challenge

Ratings of satisfaction (on the scale cited above) concern six aspects of challenge: the extent of opportunities on the job for the respondent to develop his or her own special abilities; the extent to which the work is interesting; the degree of freedom the respondent has regarding decisions about how to do his or her work; the extent to which the respondent's job gives him or her a chance to do the things he or she does best; the level of difficulty of the problems the worker is asked to solve; and the extent to which the worker can see the results of his or her work.

Satisfaction with comfort

Seven single measures of satisfaction contribute to the index of satisfaction with comfort: the convenience of travel to and from work; not being asked to do excessive amounts of work; the hours worked; freedom from conflicting demands; the amount of time to get the job done; the extent to which the respondent can forget about personal problems while at work; and the physical surroundings on the job.

Satisfaction with resources

Similar ratings of satisfaction concern three aspects of resources: the amount of help and equipment the respondent is given to get the job done; the amount of information he or she has to get the job done; and the competence of his or her supervisor.

Satisfaction with financial rewards

This index consists of ratings of satisfaction with three financial considerations: pay; job security; and fringe benefits.

Facet-specific job satisfaction

This index is a composite of the preceding five indices (i.e., satisfaction with coworkers, challenge, comfort, resources, and financial rewards) and thus is based on 21 items.

Absence of job depression

This index uses the scale--never, rarely, sometimes, often--to determine how frequently respondents experience 10 types of depressed feelings or thoughts when at work: feeling down-hearted and blue (*never*), getting tired for no reason (*never*), feeling restless and unable to keep still (*never*), finding it easy to do the things they used to do (*often*), feeling that their minds are as clear as they used to be (*often*), feeling hopeful about the future (*often*), finding it easy to make decisions (*often*), feeling more irritated than usual (*never*), still enjoying the things they used to (*often*), and feeling useful and needed (*often*).

Financial equity

The index of financial equity contains questions about whether the respondent gets more or less pay than he or she ought to compared to four groups: other people at the same employer doing similar work, other people at the same employer doing different work, other people with a different employer but with similar skills, and other people with a different employer but with a similar level of education. For each question the response alternatives are: much less than I ought to get, somewhat less than I ought to get, about the same as I ought to get, somewhat more than I ought to get, and *much more than I ought to get.*

Types of Nonwork Activities

The literature on leisure includes a number of different approaches to measuring nonwork activities. Despite substantial overlap, most studies of off-job pursuits have their own idiosyncratic list of activities, for example: Matthews and Abu-Laban (1959), 76 items; Havighurst (1961), list of all favorite leisure activities; de Grazia (1961), lists of 21 and 25 items; Converse and Robinson (1966), 27 items; Morris, Pasewark and Schultz (1972), nine items; Orthner (1975), 81 items; and London, Crandall and Fitzgibbons (1976), 30 items. The list of activities used in the present study, while fairly thorough (25 items), inadvertently omits watching television (Table 2).

Investigators have also used a variety of question formats to measure involvement in off-the-job activities: recency formats (i.e., how recently have you...), dichotomous formats (participation versus not, e.g., Matthews and Abu-Laban, 1959; Kaplan, 1960), frequency formats (e.g., Meyersohn, 1961), the time budget approach (amount of time spent on each task, e.g., de Grazia, 1961; Goldthorpe, Lockwood, Bechhofer and Platt, 1969; Willmott, 1971; Orthner, 1975), and the approach based on financial

Table 2

Recency of R's Participation in Nonwork Activities

	Activity	within last couple days	about week ago	about 2 wks ago	about a month ago	a few months ago	about a year ago	not in last year	N*
1.	When was the last time you went to the movies?	6%	7%	12%	16%	27%	8%	24%	649
2.	...went to a sports event	7	6	9	9	22	14	33	649
3.	...played some sport yourself	16	14	6	8	22	6	28	647
4.	...ate in a restaurant**	37	26	14	10	9	1	3	645
5.	...went to a bar or a nightclub**	17	14	7	12	15	6	29	647
6.	...went shopping for something besides groceries**	38	32	12	12	5	0	1	647
7.	...went to a play or concert	2	2	3	7	16	13	57	646
8.	...went to a fair, exhibit, or museum	2	2	3	6	36	19	32	648
9.	...went to a class, talk, or lecture	10	10	6	10	15	7	42	647
10.	...got in touch with relatives not living with you	56	23	10	6	4	0	1	647
11.	...got together with friends other than relatives	43	24	13	11	5	1	3	649
12.	...chatted on the phone with friends	64	16	6	5	4	1	4	648

Table 2 (continued)

Recency of R's Participation in Nonwork Activities

	Activity	within last couple days	about a week ago	about 2 weeks ago	about a month ago	a few months ago	about a year ago	not in last year	N*
13.	...made a repair or fixed something around the house	41%	19%	9%	11%	10%	1%	9%	649
14.	...sewed, embroidered, or some type of needlework**	20	9	5	8	7	3	48	624
15.	...finished reading a book	16	17	8	15	16	6	22	646
16.	...read a whole article in magazine	44	23	11	8	5	2	7	647
17.	...played cards or some other indoor game	26	20	12	11	13	3	15	648
18.	...worked on some hobby of yours	32	13	8	10	11	4	22	647
19.	...went hunting or fishing	2	3	3	5	22	11	54	648
20.	...met and talked with any people--other than those met at work--that you never met before	33	21	12	12	12	4	6	638
21.	...went out with your family	33	24	14	13	9	2	5	649
22.	...went for a pleasure trip in your car or drove to somewhere you could enjoy yourself	20	17	15	15	21	6	6	649

Table 2 (continued)

Recency of R's Participation in Nonwork Activities

Activity	a lot of times	once in a while	never	N*
23. Last summer how often did you go camping or hiking?	17%	31%	52%	649
24. ...go swimming or boating or other king of water sports	35	36	29	649
25. ...garden or work around the yard	47	29	24	649

* Excludes missing data.
** Not included in the summary index of nonwork activities.

budgets (amounts of money spent on each activity, e.g., de Grazia, 1962; Meyerson, 1969). As Table 2 indicates, the present investigation employs a recency format for most activities and a frequency format for the remainder.

Differences in lists of activities, question formats, samples, procedures for data reduction, and interpretative approaches have produced varying taxonomies of activities. London et al. (1976) have cited several examples of empirical taxonomies, ending with their own:

(1) Bishop (1970), adults, factor analysis, three factors (active/diversionary, potency, and status);

(2) Witt (1971), high school students, factor analysis, four factors (sports, outdoor-nature, adolescent-social, ant aesthetic-sophisticate);

(3) Howard (1976), high school students, factor analysis, four factors (sports, outdoor-nature, adolescent-social, and leisure detachment);

(4) McKechnie (1974), factor analysis and spherical analysis, seven factors (mechanics, crafts, intellectual, slow living, neighborhood sports, glamour sports, and fast living);

(5) Ritchie (1975), similarity ratings of leisure activities, four bipolar dimensions in his first analysis (active/passive, individual/group, simple/difficult to perform, involving/time-filling), and five types in his second analysis (active sport, relaxing entertainment, social interaction, achievement-oriented hobbies, and shopping);

(6) London et al. (1976), college students, factor analysis, three factors (sports, cultural-passive, productive-intellectual).

The six-factor taxonomy developed in the present study matches none of these earlier taxonomies exactly but does include the three factors reported by London et al. To be specific, three standard procedures for data reduction (factor analysis, cluster analysis, and correlograms) exhibited general agreement with regard to appropriate indices of off-the-job activities. The initial factor analysis, a principal components analysis with rotation to a Varimax criterion, included one item for each of the 25 activities. It produced a seven-factor solution which explained 50% of the variance but which contained certain imperfections. Some items loaded fairly evenly on several factors (e.g., went to bar or nightclub), others loaded high on none (e.g., shopped for something besides groceries), and certain factors had only one item with a substantial loading (e.g., factor five, sewed or embroidered). A second factor analysis omitted four such problematic items and thus included items on only 21 activities (Table 2). It in turn produced a six-factor solution which likewise accounted for 50% of the variance. Loadings of all items on all factors appear in Table 3. The six factors were assigned the following labels, with the items that loaded high on each factor listed in parentheses: sports & games (played some sport, went swimming or boating, went to a sports event, played indoor games, went hunting or fishing, went camping or hiking); cultural activities (went to a play or concert, went to a fair or exhibit, went to a class or lecture); intellectual activities (finished a book, read a whole magazine article); social activities (got together with friends, telephoned friends, met new people); family activities (got in touch with relatives, went out with family); and, domestic activities (did garden or yard work, made a household repair, went to the movies [scored in the opposite direction]). As noted, the other two types of data

Table 3

Factor Loadings of 21 Nonwork Activities on Six Orthogonal Dimensions
(Principal Components Analysis, Varimax Rotation)

Item	Sports & games 1	Cultural 2	Intellectual 3	Social 4	Family 5	Domestic 6	h^2 7
1. Went to the movies	0.29	-0.20	0.07	-0.17	-0.07	0.40	0.32
2. Went to a sports event	0.50	-0.22	0.05	-0.10	0.03	0.05	0.31
3. Played some sport	0.64	-0.14	0.11	-0.17	0.02	-0.01	0.47
4. Ate in a restaurant	0.21	-0.14	-0.04	-0.30	-0.23	0.08	0.22
5. Went to a play or concert	0.04	-0.70	0.12	-0.11	-0.03	0.01	0.52
6. Went to a fair or exhibit	0.22	-0.37	0.14	-0.09	-0.02	0.08	0.22
7. Went to a class or lecture	0.06	-0.34	0.29	-0.03	-0.14	-0.05	0.23
8. Got in touch with relatives	-0.01	-0.03	0.06	-0.11	-0.35	0.02	0.14
9. Got together with friends	0.22	-0.05	0.08	-0.69	-0.17	0.05	0.57
10. Telephoned friends	0.01	-0.09	0.20	-0.52	-0.01	0.01	0.32
11. Made a household repair	0.22	-0.01	0.12	-0.08	-0.07	-0.40	0.23
12. Finished a book	0.04	-0.19	0.62	-0.12	0.01	-0.04	0.44
13. Read a whole magazine article	0.20	-0.10	0.50	-0.18	-0.11	0.03	0.34
14. Played indoor games	0.44	0.06	0.01	-0.23	-0.08	-0.05	0.26
15. Worked on hobby	0.29	-0.09	0.15	-0.14	-0.15	-0.22	0.20

Table 3 (continued)

Factor Loadings of 21 Nonwork Activities on Six Orthogonal Dimensions
(Principal Components Analysis, Varimax Rotation)

Item	Sports & Games 1	Cultural 2	Intellectual 3	Social 4	Family 5	Domestic 6	h^2 7
16. Went hunting or fishing	0.36	0.05	0.02	0.11	-0.25	-0.20	0.25
17. Met new people	0.16	-0.16	0.14	-0.30	-0.24	-0.08	0.22
18. Went out with family	0.07	0.02	-0.07	-0.11	-0.36	-0.24	0.20
19. Went camping or hiking	0.31	-0.15	0.22	0.02	-0.34	-0.07	0.29
20. Went swimming or boating	0.55	-0.14	0.22	0.02	-0.41	0.04	0.54
21. Did garden or yard work	-0.02	-0.01	-0.02	0.05	-0.03	-0.51	0.26

reduction, cluster analysis and correlograms, produced quite similar patterns of items and suggested no new plausible groupings.

The second factor analysis thus generated six indices of activities. Quite simply, scores on items loading high in the factor analysis were summed for all six activity factors: sports & games, cultural, intellectual, social, family, and domestic. In addition, a seventh index, political activities was added by summing scores on four separate political items (Table 4). All four indicators of political activity intercorrelated highly and positively (i.e., above 0.40); interest in 1972 campaign; being registered to vote; having voted or planning to vote in 1972; and frequency of voting in previous presidential campaigns.

Two items in the interview tapped involvement with friends and neighbors: number of close friends not counting relatives and neighbors; and number of neighbors the respondent knew well enough to visit with (Table 5). Because of the low (positive) correlation between those two items ($r=0.17$, $df=646$, $p < 0.01$), the analysis treats them separately and does not combine them into a single index.

Degree of Nonwork Involvement

Several attempts at data reduction failed to provide meaningful indices of organizational involvement measured in terms of organizational membership. Interviewers asked respondents which of 17 organizations they belonged to, and whether or not they belonged to any other organizations. Table 6 lists these organizations, as well as their univariate distributions. The analysis of organizational membership included the three earlier types of data reduction: factor analysis, cluster analysis, and correlograms. None of the three techniques, however,

Table 4

Frequency Distribution on Items on Political Involvement

Interest in 1972 political campaign (n=649)*	
Very interested	36%
Somewhat interested	42
not much interested	22

Currently registered to vote (n=649)*	
Yes	74%
No	26

Voted/plan to vote in November 1972[a] (n=485)*	
Yes	84%
No	16

Frequency of voting in previous presidential elections[b] (n=604)*	
All	48%
Most	18
A few	15
None	19

---

* Excludes missing data.

[a] Excludes those not registered to vote.

[b] Refers only to those elections for which R was old enough to vote.

Table 5

Frequency Distribution on Items Concerning Friends and Neighbors

Number of close friends[a] (n=648)*

0-2 friends	21%
3-4	20
5-6	21
7-12	22
14 and above	16

Number of neighbors R knows well enough to visit (n=648)*

0 neighbors	22%
1-2	29
3-5	26
6 and above	23

---

* Excludes missing data.

[a] Not counting relatives.

Table 6

Percent R's Currently Belonging to Clubs, Groups and Organizations

Item	Percent R's belonging to organization	N*
1. Sports club or team	21%	648
2. Social or card playing group	15	648
3. Church or synogogue	47	649
4. Church- or synogogue-connected group or association	13	649
5. Lodge, Fraternity, sorority, or veteran's organization	13	648
6. Labor union**	19	648
7. A cooperative	8	647
8. Nationality, ethnic, or racial association	3	648
9. Professional association**	13	648
10. Social welfare or charity group	3	647
11. Parent-teacher association**	10	649
12. Youth groups, like being a Girl Scout leader or a little league coach**	8	648
13. Country club	3	647
14. Community center	2	647
15. Discussion group	6	648
16. Neighborhood or community improvement group	5	647
17. Political club or organization	3	647
18. Other clubs, groups, organizations **	14	647

* Excludes missing data.
** Not included in the index of number of organizational memberships.

generated meaningful groupings of organizations, hence the decision to abandon indices of organizational types. Instead the analysis focused on the (absolute) number of organizations to which each respondent belonged, later revised to the (absolute) number of organizations open to all respondents, no matter what their demographic characteristics (indicated in Table 6), with a maximum of 13.

Two other summary indices of nonwork involvement combined items and indices already described. The index of overall nonwork activity comprised the sum of scores on the six activity factors from the principal components analysis (sports & games, cultural, intellectual, social, family, and domestic). The index of overall nonwork involvement, similarly, represented the sum of the same six activity factors plus the index of political activities, the two items on friends and neighbors, and the number of organizational memberships.

Resources for Nonwork Activities

The study included measures of four basic resources needed for off-the-job participation: time, energy, health, and money. The procedures for selecting measures required that each measure of a resource meet two criteria: at least as many positive as negative intercorrelations with other measures of the same resource; and significant positive correlations with at least one of the nonwork activities. The following 11 measures of the four resources survived these two tests:

Time - degree to which the worker says he or she controls the overtime hours he or she works

- worker's reported satisfaction with the hours he or she works

- presence and severity of problems related to worker's hours, schedule, or overtime*

Energy - self-rating by worker on a ladder of pep-and-energy

Health - reported number of somatic complaints experienced in the past year, from a list of 22 items*
- reports of job-related illnesses and injuries within the past year*
- reported number of common illnesses and injuries suffered in the past year, with a maximum score of 14*
- self-rating on a ladder of general health
- number of days the respondent reports he or she did not go to work in the past month because he or she did not feel like going*
- interviewer's report that the respondent had either a speech defect or a physical disfigurement*

Money - self-report of personal annual income from job

Analysis Plan

The first step in the analysis was to remove the effects of several basic resources on participation in nonwork roles. Measures of the four major resources--time, energy, health, and money--were, therefore, correlated with measures of types of nonwork activities. For each index of nonwork activities, all the measures of resources that met the two criteria for selection (i.e., correlated positively with other measures of the same resource; and correlated positively and significantly with the index) were included as predictors of the index in a multiple regression. The effects of the measures of resources on each index of nonwork activity were then removed statistically via the regression, leaving residualized scores for all respondents. Subsequent analyses used both nonresidualized and residualized scores on nonwork participation, the latter providing a statistical control on the availability of several important resources.

---

*These measures of resources are scored in the reversed direction.

As a second step in analysis, the measures of involvement in the work role (degree of involvement, types of activities, and subjective reactions) were correlated with both nonresidualized and residualized measures of nonwork involvement (degree of involvement, types of activities), and the correlations were assigned to the six relevant cells in Table 1 (one, two, four, five, seven, and eight). The correlational strategy was then repeated for demographic subgroups of workers defined in terms of sex, marital status (married versus nonmarried), age (less than 32, 32 and over), and education (less than a high school diploma, high school diploma, more than a high school diploma).

Results

The results presented in this section address several different issues. The initial tables display univariate distributions for the single items measuring nonwork activities. Next are the tables depicting correlations between measures of resources and measures of nonwork activities. The major tables in the present analysis, those correlating measures of involvement in work and nonwork, appear third. Cells one, two, four, five, seven and eight in Table 1 each have at least one table depicting the correlations between work and nonwork for the total sample and another table summarizing parallel correlations for various demographic subsamples. All such tables present correlations based on both nonresidualized and residualized scores.

Univariate Distributions of Measures of Nonwork Activities

Table 2 indicates how recently respondents had engaged in 25 activities. Those activities undertaken most recently (and hence most frequently) concern contact with relatives and friends, plus such basic activities as shopping, reading an article, and eating at a restaurant. Those engaged in least recently include cultural activities as well as those more seasonal and sex-specific such as hunting and fishing. For the final three items which use a frequency format, respondents reported more frequent participation in water sports and yard work than in camping and hiking.

As regards the univariate distributions on measures of political activities shown in Table 4, clear majorities of respondents described themselves as somewhat or very interested in the 1972 political campaign, reported that they were registered to vote, claimed that they had voted or intended to vote in the 1972 election, and observed that they had voted in most or all of those previous presidential elections for which they had been old enough.

Univariate distributions for items on friends and neighbors appear in Table 5. One fifth of the sample reported having no more than two close friends, not counting relatives or neighbors, another fifth mentioned three or four close friends, a further fifth said five or six, yet another fifth cited seven to 12, and the final fifth claimed more than 12 close friends. Almost a quarter of all respondents claimed not to know any neighbors well enough to visit with. Somewhat over a quarter had one or two neighbors they knew well enough to visit. A third quarter of the respondents cited three to five such neighbors, and the final quarter cited more than five.

According to Table 7, workers in the sample belonged to few of the 13 associations considered open to all. One third had no affiliation with any organization, roughly another 30 percent belonged to only one association, and only about 10 percent of the sample belonged to four or more organizations.

Table 6 reports the percent of respondents belonging to each of 17 specific organizations. Church or synogogue leads the list with nearly half the respondents having reported that they belonged to one or the other. Roughly 20 percent belonged to a sports club and a similar percentage belonged to a labor union. Trailing the list, with only three percent of respondents having reported membership are nationality/ethnic/racial associations, social welfare/charity groups, country clubs, community centers, and political organizations.

Correlations between Measures of Resources and Nonwork Involvement

Before the positive and negative hypotheses may be tested on residualized as well as nonresidualized measures of nonwork involvement, the relationship between resources and nonwork involvement must be clarified. Table 8 presents the correlations between 11 measures of four resources (time, energy, money, and health) and 12 measures of involvement in nonwork activities (nine measures of types of involvement and three summary indices

Table 7

Number of Organizations to which R Belongs[a] (n=646)*

Zero	33%
1	29
2	19
3	10
4	5
5 or more	4

---

* Excludes missing data

[a] Includes only 13 specific organizations accessible to most workers, listed in Table 6.

Table 8

Correlations between Measures of Resources and Measures of Nonwork Involvement

Resources	Nonwork Involvement								
	Sports & games	Cultural	Intellectual	Social	Family	Domestic	Political activities	Number close friends	# Neighbors know well enough to visit
Time									
Control over overtime	.06 (618)	.18** (618)	.19** (614)	.09* (606)	.05 (616)	.02 (618)	.22** (618)	-.03 (617)	-.05 (618)
Satisfaction with working hours	-.10* (647)	.03 (647)	.02 (643)	-.04 (635)	-.07 (645)	-.01 (647)	.11** (647)	-.05 (646)	.03 (646)
Severity of hours problems (scored in the direction of low severity)	-.05 (637)	-.06 (637)	.01 (633)	-.02 (626)	-.02 (635)	.08 (637)	.05 (637)	.02 (636)	.08 (636)
Energy									
Pep-and-energy ladder	.12** (648)	.17** (648)	.11** (644)	.10* (637)	.07 (646)	.05 (648)	.11** (648)	.05 (647)	.06 (647)
Money									
Personal annual income	.17** (645)	.15** (645)	.13** (641)	.00 (634)	.06 (643)	.19** (645)	.15** (645)	.10* (644)	.05 (644)
Health									
Number of psychosomatic illnesses (scored in direction of low number)	.04 (643)	.12** (643)	.01 (639)	.00 (631)	.06 (641)	.04 (643)	.05 (643)	.04 (642)	.07 (642)
Occurrence of job-related illness or injury (scored in the direction of no illness or injury)	.00 (649)	.05 (649)	.02 (645)	-.03 (637)	.01 (647)	.07 (649)	.18** (649)	-.06 (648)	.06 (648)
Number of illnesses (scored in the direction of low number)	.06 (649)	.06 (649)	-.03 (645)	.06 (637)	.10** (647)	-.01 (649)	.00 (649)	-.05 (648)	.03 (648)
Health ladder	.17** (648)	.20** (648)	.10* (644)	.09* (637)	.09* (646)	.03 (648)	.13** (648)	-.05 (647)	.06 (647)
Number of days not feel like going to work (scored in direction of few days)	-.05 (647)	.07 (647)	.09* (643)	.08* (635)	.04 (645)	.15** (647)	.12** (647)	.02 (646)	.05 (646)
Presence of speech defect or physical disfigurement (scored in direction of no defect/disfigurement)	.13** (638)	.12** (638)	.09* (634)	.07 (626)	.05 (636)	-.03 (638)	.01 (638)	-.04 (637)	.04 (637)

Table 8 (continued)

Correlations between Measures of Resources and Measures of Nonwork Involvement

Resources	Nonwork Involvement		
	Memberships in selected organizations	Overall nonwork activity	Overall nonwork involvement
Time			
Control over overtime	.16** (616)	.17** (601)	.20** (597)
Satisfaction with working hours	.06 (645)	-.03 (630)	.02 (625)
Severity of hours problems (scored in the direction of low severity)	.12** (635)	.01 (621)	.05 (616)
Energy			
Pep-and-energy ladder	.16** (646)	.17** (632)	.19** (627)
Money			
Personal annual income	.23** (643)	.22** (629)	.25** (624)
Health			
Number of psychosomatic illnesses (scored in direction of low number)	.12** (641)	.08* (626)	.11** (621)
Occurrence of job-related illness or injury (scored in the direction of no illness or injury)	.10** (647)	.03 (632)	.09* (621)
Number of illnesses (scored in the direction of low number)	.02 (647)	.05 (632)	.04 (627)
Health ladder	.13** (646)	.19** (632)	.19** (627)
Number of days not feel like going to work (scored in direction of few days)	.14** (645)	.11** (630)	.14** (625)
Presence of speech defect or physical disfigurement (scored in direction of no defect/disfigurement)	.02 (636)	.15** (621)	.11* (616)

of degree of involvement). The table permits a number of comparisons among measures of resources.

For each resource, most of the correlations follow the predicted positive pattern: 24 positive correlations out of 36 for time; 12 out of 12 for energy; 12 out of 12 for money; and 63 out of 72 for health. The statistically significant correlations follow the positive pattern even more consistently: nine significantly positive correlations out of 10 for time; eight out of eight for energy; nine out of nine for money; and 29 out of 29 for health. Money emerges as the resource significantly and positively associated with the widest array of nonwork activities (nine significantly positive correlations out of 12 correlations; median $r=0.15$, $df=643$, $p < 0.01$), closely followed by energy (eight out of 12; median $r=0.11$, $df=642$, $p < 0.01$). Next comes health (29 out of 72; median $r=0.06$, $df=635$, n.s.); and time shows the weakest pattern (nine out of 36; median $r=0.03$, $df=644$, n.s.).

At the level of single measures, those most closely associated with off-the-job activities include a measure of money (personal income, nine significantly positive correlations out of 12, median r as above for money), a measure of health (health ladder, nine out of 12, median $r=0.13$, $df=644$, $p < 0.01$), and a measure of energy (pep and energy ladder, eight out of 12, median r as above for energy). Three single measures turn up only one significantly positive correlation: satisfaction with working hours (median $r=0.02$, $df=623$, n.s.), severity of problem(s) regarding working hours (median $r=0.02$, $df=634$, n.s.), and number of illnesses and injuries suffered in the past year (median $r=0.05$, $df=630$, n.s.). Among single measures, only one generates a statistically significant negative correlation: satisfaction with hours appears inversely related to participation in sports & games.

Table 8 also permits comparisons of correlations between resources and nonwork involvement among the measures of nonwork activities. As regards consistent patterns of positive correlations, political activities, memberships in organizations, and overall nonwork involvement score highest (all have 11 positive correlations with resources out of 11), followed by cultural activities, intellectual activities, number of neighbors friendly, and overall nonwork activity (10 out of 11). Scoring lowest is the item concerning number of close friends (five out of 11). With the criterion raised to number of significantly positive correlations, memberships in organizations and overall nonwork involvement both register the top score (eight out of 11), followed by political activities and overall nonwork activity (both seven out of 11). Number of neighbors friendly scores lowest (zero significantly positive correlations out of 11), along with number of close friends (one out of 11), and family activities and domestic activities (both two out of 11). With size of median correlation as the criterion, five measures of nonwork activities show the strongest relationship: cultural activities, political activities, number of organizational memberships, overall nonwork activities, and overall nonwork involvement (median r's=0.11-0.12, $df \geq 614$, $p < 0.01$). Five other measures produce the lowest median correlations: sports & games, social activities, family activities, domestic activities, number of close friends, and number of neighbors known well enough to visit (median r's < 0.07, $df \leq 642$, n.s.).

The table thus sustains several inferences. At a methodological level, the three summary measures of degree of involvement in nonwork relate, as expected, more strongly to the measures of resources than do the nine measures of types of nonwork involvement. On a substantive level, public-oriented activities off the job (e.g., political activities and organizational memberships) require more resources than do home-based activities (e.g., family activities, domestic activities, having close friends, and having friendly neighbors). Certain more specific inferences based on Table 8 require a simultaneous examination of types of resources and types of nonwork involvement. Control over overtime correlates with the more optional or discretionary types of nonwork involvement (e.g., cultural, intellectual, social, and political activities, and organizational memberships). Although measures of energy and health correlate strongly with many of the indices of nonwork, energy and health appear less necessary, according to the table, for interpersonal activities involving family, friends, and neighbors. To a reduced extent, money also relates less strongly to informal interpersonal activities. Put another way, money fails to correlate strongly with involvement in relatively inexpensive, interpersonal activities.

Table 8 indicates, finally, those resources whose effects need to be removed statistically from the indices of nonwork involvement. Quite simply, for each nonwork item or index the list of predictors whose effects are to be removed via multiple regression consists of those measures of resources that correlate positively and significantly with that item or index. Each measure of nonwork involvement thus becomes the dependent variable for one regression, its label appearing at the top of one column in Table 8. Using asterisks to indicate significance, the remainder of the column designates those measures of resources to be included as predictors in the multiple regression. Thus the multiple regression with sports & games as the dependent variable has four predictors: pep and energy ladder, personal income, health ladder, and presence of speech defect or physical disfigurement. The multiple regression with cultural activities as the dependent variable, similarly, has six predictors: control over overtime, pep and energy ladder, annual income, number of somatic complaints, health ladder, presence of speech defect or physical disfigurement, and so on. One of the measures of nonwork involvement, number of friendly neighbors, correlates significantly with none of the resource predictors and hence requires no multiple regression.

Results of the 11 ensuing multiple regressions appear in Table 9. All regressions generate statistically significant multiple correlations. As expected, regressions with more predictors result in higher multiple correlations. The regression procedure also provides residualized scores for all respondents on the 11 measures of nonwork involvement from which the effects of the relevant resources were removed.

Testing Hypotheses Relating to the Cells in Table 1

Cell one concerns the correlations between 12 measures of degree of work involvement (three objective and nine subjective) and three measures of degree of nonwork involvement. Tables 10 and 11 present these correlations for both nonresidualized and residualized measures of nonwork involvement and finds support for the positive hypothesis for the subjective (or attitudinal) but not for the objective measures of work involvement. The objective measures, which tap time devoted to the work role, are mostly

Table 9

Multiple Regressions of Measures of Resources on Measures of Nonwork Involvement

Index Off-the-job Involvement	Resource Predictors	Multiple R	N	Signif.
Sports & games	Pep-and-energy ladder Personal annual income Health ladder Presence of speech defect or physical disfigurement	.26	633	.01
Cultural	Control over overtime Pep-and-energy ladder Personal annual income Number of psychosomatic illnesses Health ladder Presence of speech defect or physical disfigurement	.32	597	.01
Intellectual	Control over overtime Pep-and-energy ladder Personal annual income Health ladder Number of days not feel like going to work Presence of speech defect or physical disfigurement	.24	597	.01
Social	Control over overtime Pep-and-energy ladder Health ladder Number of days not feel like going to work	.14	604	.02
Family	Number of illnesses Health ladder	.12	646	.02
Domestic	Personal annual income Number of days not feel like going to work	.24	643	.01
Political activities	Control over overtime Satisfaction with working hours Pep-and-energy ladder Personal annual income Occurrence of job-related illness or injury Health ladder Number of days not feel like going to work	.31	609	.01

Table 9 (continued)

Multiple Regressions of Measures of Resources on Measures of Nonwork Involvement

Index Off-the-job Involvement	Resource Predictors	Multiple R	N	Signif.
Number of close friends	Personal annual income	.10	644	.02
Number neighbors know well enough to visit	None			
Memberships in selected organizations	Control over overtime Severity of hours problems Pep-and-energy ladder Personal annual income Number of psychosomatic illnesses Occurrence of job-related illness or injury Health ladder Number of days not feel like going to work	.32	591	.01
Overall nonwork activity	Control over overtime Pep-and-energy ladder Personal annual income Number of psychosomatic illnesses Health ladder Number of days not feel like going to work Presence of speech defect or physical disfigurement	.35	580	.01
Overall nonwork involvement	Control over overtime Pep-and-energy ladder Personal annual income Number of psychosomatic illnesses Occurrence of job-related illness or injury Health ladder Number of days not feel like going to work Presence of speech defect or physical disfigurement	.37	576	.01

Table 10 (Cell 1)

Correlations between Measures of Degree of Work Involvement and Degree of Nonwork Involvement for Total Sample

Work Involvement	Nonwork Involvement					
	Memberships in selected organizations		Overall nonwork activity		Overall nonwork involvement	
	non-res.	res.	non-res.	res.	non-res.	res.
Objective Measures						
Hours worked per week	.04 (643)	-.01 (587)	.03 (629)	-.05 (577)	.07 (624)	-.02 (573)
Have second job	.02 (646)	.01 (590)	.03 (631)	.04 (579)	.04 (626)	.05 (575)
Hours worked on second job	-.04 (51)	-.05 (47)	-.19 (51)	-.11 (47)	-.20 (51)	-.14 (47)
Subjective Measures						
Do my best at work	-.07 (643)	-.06 (587)	-.08* (628)	-.09* (576)	-.05 (623)	-.04 (572)
Work my hardest at work	.04 (644)	.02 (588)	.07 (629)	.01 (577)	.11** (624)	.06 (573)
Would work even if enough money	.08* (645)	.03 (589)	.08* (630)	.04 (578)	.11* (625)	.05 (574)
Time drags at work	.23** (646)	.13** (590)	.21** (631)	.10* (579)	.25* (626)	.14** (575)

Table 10 (Cell 1 continued)

Correlations between Measures of Degree of Work Involvement and Degree of Nonwork Involvement for Total Sample

Work Involvement	Nonwork Involvement: Memberships in selected organizations non-res.	res.	Overall nonwork activity non-res.	res.	Overall nonwork involvement non-res.	res.
Subjective Measures						
Ideal job if free to choose	.04 (601)	-.03 (548)	-.02 (590)	-.07 (541)	-.01 (585)	-.07 (537)
Job skills useful in future	.11* (644)	.02 (589)	.12** (631)	.03 (579)	.13** (626)	.03 (575)
Degree absorbed in job	.19** (644)	.09* (589)	.10* (630)	.01 (578)	.17** (625)	.08 (574)
Work harder than coworkers	.10** (644)	.08* (587)	.12** (627)	.10* (577)	.12** (622)	.10* (573)
Frequency of extra work	.12** (645)	.07 (590)	.15** (631)	.11** (579)	.17** (626)	.12** (575)

*p<0.05
**p<0.01

Table 11 (Cell 1)

Correlations between Measures of Degree of Work Involvement and Degree of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations		Significant Correlations		Median Correlations	
		No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation	N
OBJECTIVE MEASURES							
Total Sample	nonres.	6	9	0	0	.03	629
	res.	3	9	0	0	-.02	573
Males	nonres.	5	6	0	0	.02	319
	res.	3	6	0	0	.02	301
Females	nonres.	3	6	0	0	.00	313
	res.	2	6	0	0	-.01	289
Males, married	nonres.	3	6	0	0	.01	259
	res.	3	6	0	0	.02	241
Females, married	nonres.	0	6	0	0	-.03	166
	res.	1	6	0	0	-.05	155
Males, not married	nonres.	5	6	0	0	.09	60
	res.	5	6	0	0	.16	56
Females not married	nonres.	6	6	0	0	.03	142
	res.	6	6	0	0	.03	133
Males, under 32 years	nonres.	4	6	1	1	.10	148
	res.	3	6	2	2	.14	135
Females, under 32 years	nonres.	4	6	0	0	.01	163
	res.	3	6	0	0	.02	149
Males, 32 years or over	nonres.	2	6	0	0	-.04	171
	res.	2	6	0	0	-.05	163
Females, 32 years or over	nonres.	3	6	0	0	.01	151
	res.	1	6	0	0	-.03	13[illegible]
Males, less than 12th grade	nonres.	4	6	0	0	.02	86
	res.	1	6	0	0	-.02	82
Females, less than 12th grade	nonres.	3	6	0	0	.00	80
	res.	3	6	0	0	.10	68
Males, 12th grade	nonres.	4	6	0	0	.07	130
	res.	4	6	0	0	.05	122
Females, 12th grade	nonres	1	6	0	0	-.02	135
	res.	1	6	0	0	-.04	131
Males, more than 12th grade	nonres.	3	6	0	0	.06	103
	res.	2	6	0	1	-.02	97
Females, more than 12th grade	nonres.	6	6	0	0	.07	96
	res.	5	6	0	0	.03	81

Table 11 (Cell 1 continued)

Correlations between Measures of Degree of Work Involvement and Degree of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
SUBJECTIVE MEASURES							
Total Sample	nonres.	22	27	19	20	.11**	624
	res.	21	27	9	10	.04	578
Males	nonres.	23	27	12	15	.09	317
	res.	14	27	2	5	.01	297
Females	nonres.	24	27	10	10	.08	315
	res.	22	27	5	5	.04	288
Males, married	nonres.	24	27	11	13	.09	257
	res.	15	27	2	3	.03	246
Females, married	nonres.	22	27	6	6	.11	172
	res.	22	27	4	4	.07	155
Males, not married	nonres.	15	27	2	4	.05	61
	res.	17	27	2	4	.03	55
Females, not married	nonres.	21	27	6	6	.06	142
	res.	15	27	2	4	.03	131
Males, under 32 years	nonres.	23	27	4	4	.11	146
	res.	17	27	2	5	.03	135
Females, under 32 years	nonres.	24	27	13	13	.15	157
	res.	22	27	6	7	.08	145
Males, 32 years or over	nonres.	19	27	5	8	.09	178
	res.	16	27	2	2	.03	163
Females, 32 years or over	nonres.	20	27	3	3	.04	152
	res.	15	27	2	2	.01	124
Males, less than 12th grade	nonres.	23	27	5	5	.08	85
	res.	18	27	3	3	.07	79
Females, less than 12th grade	nonres.	23	27	6	6	.08	75
	res.	22	27	4	4	.07	65
Males, 12th grade	nonres.	18	27	3	6	.06	126
	res.	13	27	0	3	-.01	119
Females, 12th grade	nonres.	23	27	8	8	.11	135
	res.	17	27	3	3	.08	130
Males, more than 12th grade	nonres.	26	27	8	8	.15	103
	res.	17	27	0	1	.04	97
Females, more than 12th grade	nonres.	16	27	2	2	.01	96
	res.	4	27	0	0	-.05	81

unrelated to either the nonresidualized or residualized measures of degree of nonwork involvement. To be specific, six of the nine correlations are positive for the nonresidualized scores (three out of nine for the residualized scores), none of the correlations is significant for either the nonresidualized or residualized scores, and the median correlations are trivial for both types of scores ($r=0.03$, $df=627$, n.s.; $r=-0.02$, $df=571$, n.s., respectively).

Comparisons of demographic subgroups, based on only two objective measures because of small sample sizes for the third measure (viz. hours spent on second job), reveal only limited variations in the correlational pattern. They show for both nonresidualized and residualized scores a somewhat greater number of positive and significantly positive correlations among males under 32.

As noted, the subjective measures of degree of work involvement provide substantial empirical support for the positive hypothesis: mainly positive correlations (22 of 27 for the nonresidualized scores, 21 of 27 for the residualized scores), mainly positives among the significant correlations (19 of 20 for the nonresidualized scores, nine of 10 for the residualized scores), and a substantial median correlation for the nonresidualized scores ($r=0.11$, $df=622$, $p < 0.01$) but not for the residualized scores ($r=0.04$, $df=576$, n.s.). Although the correlational patterns do not differ noticeably among the three measures of nonwork involvement, the various measures of work involvement do make a difference. For both nonresidualized and residualized scores, measures of feeling involved in the job (time drags at work, degree absorbed in job) and on-the-job effort (work harder than coworkers, frequency of extra work) bear the strongest positive relationship with degree of nonwork involvement. Variables least consistent with the positive pattern are: do my best at work, and ideal job if free to choose.

The differences among demographic subgroups based on subjective measures prove somewhat confusing. To take one example, for nonresidualized scores, males and females scarcely differ. Yet, for residualized scores, females are consistently more in line with the positive hypothesis than are males. Demographic subgroups in conformity with the positive hypothesis include females under 32, whereas those most noticeably at odds with the positive hypothesis include females over 31, males with a high school diploma, nonmarried males, and females with more than a high school diploma.

Table 12 presents the data relevant to cell two, namely, the correlations between the 12 measures of degree of work involvement (three objective and nine subjective) and the nine measures of types of nonwork activities. As noted, one of the nine variables, number of friendly neighbors, proved unrelated to any of the measures of resources and thus generated no separate set of residualized scores. For the other eight measures of nonwork activities, correlations based on residualized scores accompany those based on nonresidualized scores.

The overall correlational pattern in Tables 12 and 13 suggests at best a weak positive relationship between degree of work involvement and different types of nonwork activities. For nonresidualized and residualized scores, the objective measures of work involvement generate a much higher proportion of

Table 12 (Cell 2)

Correlations between Measures of Degree of Work Involvement and Types of Nonwork Involvement for Total Sample

Work Involvement	Nonwork Involvement											
	Sports & games		Cultural		Intellectual		Social		Family		Domestic	
	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.
Objective Measures												
Hours worked per week	.08* (646)	.01 (630)	-.02 (646)	-.08 (594)	-.02 (642)	-.06 (594)	-.14** (634)	-.16** (601)	.10** (644)	.10* (643)	.11** (646)	.05 (640)
Have second job	.06 (648)	.05 (632)	.08* (648)	.10* (596)	.01 (644)	.02 (596)	.06 (636)	.05 (603)	-.06 (646)	-.06 (645)	-.05 (648)	-.06 (642)
Hours worked on second job	-.31 (51)	-.28 (49)	-.17 (51)	-.13 (47)	-.23 (51)	-.23 (47)	-.04 (51)	-.03 (48)	.12 (51)	.12 (51)	.01 (51)	.03 (50)
Subjective Measures												
Do my best at work	-.05 (645)	-.06 (629)	-.04 (645)	-.06 (593)	-.21** (641)	-.21** (593)	-.09* (633)	-.09* (600)	.04 (643)	.04 (642)	.07 (645)	.07 (639)
Work my hardest at work	.02 (646)	-.01 (630)	.02 (646)	.00 (594)	-.05 (642)	-.08 (594)	-.00 (634)	-.02 (601)	.05 (644)	.06 (643)	.13** (646)	.10* (640)
Would work even if enough money	.12** (647)	.08* (631)	.07 (647)	.03 (595)	.09* (643)	.05 (595)	.01 (635)	-.01 (602)	.02 (645)	.01 (644)	-.07 (647)	-.13** (641)
Time drags at work	.06 (648)	-.01 (632)	.15** (648)	.06 (596)	.13** (644)	.07 (596)	.08* (636)	.03 (603)	.10** (646)	.09* (645)	.15** (648)	.08* (642)
Ideal job if free to choose	-.06 (604)	-.07 (589)	.03 (604)	.01 (555)	-.03 (601)	-.07 (556)	-.09* (593)	-.11** (562)	-.[illegible]0 (603)	-.01 (602)	.06 (604)	.01 (598)
Job skills useful in future	.04 (647)	-.02 (632)	.14** (647)	.08 (596)	.07 (643)	-.03 (596)	.02 (636)	.00 (603)	.05 (645)	.04 (645)	.08* (647)	.03 (641)
Degree absorbed in job	-.08* (646)	-.13** (631)	.11** (646)	.05 (595)	.02 (642)	-.03 (595)	-.01 (635)	-.05 (602)	.07 (644)	.06 (644)	.21** (646)	.14** (640)
Work harder than coworkers	.10* (643)	.08* (627)	.06 (643)	.06 (593)	.06 (639)	.06 (593)	.00 (632)	-.00 (601)	.05 (642)	.06 (641)	.06 (643)	.03 (637)
Frequency of extra work	.06 (647)	.04 (632)	.11** (647)	.09* (596)	.09* (643)	.06 (596)	.03 (636)	-.00 (603)	.05 (645)	.04 (645)	.11** (647)	.06 (641)

Table 12 (Cell 2 continued)

Correlations between Measures of Degree of Work Involvement and Types of Nonwork Involvement for Total Sample

Work Involvement	Nonwork Involvement				
	Political activities		# Close friends		# Neighbors know well enough to visit[a]
	non-res.	res.	non-res.	res.	nonresidual
Objective Measures					
Hours worked per week	.00 (646)	-.03 (607)	.18** (645)	.13** (641)	.08* (645)
Have second job	.02 (648)	.02 (608)	.05 (647)	.06 (643)	.01 (647)
Hours worked on second job	-.13 (51)	-.06 (47)	.12 (51)	.13 (50)	-.00 (51)
Subjective Measures					
Do my best at work	-.01 (645)	-.01 (606)	.09* (644)	.09* (640)	-.04 (644)
Work my hardest at work	.03 (646)	.01 (607)	.10* (645)	.09* (641)	.04 (645)
Would work even if enough money	.05 (647)	-.00 (607)	.07 (646)	.05 (642)	.05 (646)
Time drags at work	.18** (648)	.07 (608)	.09* (647)	.06 (643)	.10** (647)
Ideal job if free to choose	.06 (604)	-.00 (567)	.03 (603)	.00 (599)	-.04 (603)
Job skills useful in future	.09* (647)	.01 (608)	-.02 (646)	-.02 (642)	-.02 (646)
Degree absorbed in job	.17** (646)	.09* (607)	.08* (645)	.04 (641)	.05 (645)
Work harder than coworkers	.09* (643)	.09* (605)	.07 (642)	.05 (638)	-.01 (642)
Frequency of extra work	.16** (647)	.11* (608)	.12** (646)	.08* (642)	-.00 (646)

---

[a] For this cell and for cells five and eight, in counts of correlations based on the item "Number of neighbors know well enough to visit", residualized scores are the same as nonresidualized scores.

Table 13 (Cell 2)

Correlations between Measures of Degree of Work Involvement and Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
			OBJECTIVE MEASURES				
Total Sample	nonres.	16	27	6	7	.01	644
	res.	15	27	4	5	.01	630
Males	nonres.	13	18	2	3	.04	329
	res.	10	18	2	3	.02	307
Females	nonres.	11	18	1	1	.03	318
	res.	11	18	2	2	.03	297
Males, married	nonres.	12	18	0	1	.03	269
	res.	9	18	0	1	.02	265
Females, married	nonres.	10	18	1	2	.03	174
	res.	9	18	0	2	.00	158
Males, not married	nonres.	12	18	0	1	.12	61
	res.	12	18	1	2	.12	59
Females, not married	nonres.	12	18	2	2	.02	143
	res.	12	18	2	2	.02	138
Males, under 32 years	nonres.	13	18	1	3	.03	151
	res.	10	18	1	3	.02	151
Females, under 32 years	nonres.	11	18	2	2	.02	165
	res.	12	18	2	2	.02	164
Males, 32 years or over	nonres.	10	18	1	2	.01	180
	res.	8	18	1	2	-.01	177
Females, 32 years or over	nonres.	6	18	1	2	-.03	151
	res.	7	18	0	1	-.02	136
Males, less than 12th grade	nonres.	8	18	0	0	-.01	93
	res.	7	18	0	0	-.03	92
Females, less than 12th grade	nonres.	8	18	2	2	-.02	81
	res.	8	18	2	2	-.01	80
Males, 12th grade	nonres.	11	18	1	2	.05	129
	res.	11	18	2	3	.04	122
Females, 12th grade	nonres.	9	18	0	2	.02	137
	res.	10	18	0	2	.02	138
Males, more than 12th grade	nonres.	12	18	2	3	.08	106
	res.	8	18	1	2	-.01	101
Females, more than 12th grade	nonres.	12	18	1	1	.04	97
	res.	12	18	2	2	.04	97

Table 13 (Cell 2 continued)

Correlations between Measures of Degree of Work Involvement and Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations		Significant Correlations		Median Correlations	
		No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation	N
SUBJECTIVE MEASURES							
Total Sample	nonres.	61	81	27	31	.05	642
	res.	51	81	13	18	.03	595
Males	nonres.	56	81	18	22	.04	330
	res.	44	81	7	13	.01	329
Females	nonres.	61	81	19	23	.06	315
	res.	55	81	6	10	.04	290
Males, married	nonres.	59	81	16	21	.04	262
	res.	42	81	4	9	.01	249
Females, married	nonres.	66	81	14	16	.06	172
	res.	59	81	6	7	.04	174
Males, not married	nonres.	39	81	4	9	-.01	60
	res.	39	81	5	10	-.01	60
Females, not married	nonres.	55	81	12	13	.05	143
	res.	47	81	8	13	.02	133
Males, under 32 years	nonres.	66	81	9	12	.04	150
	res.	42	81	3	7	.00	149
Females, under 32 years	nonres.	65	81	12	13	.06	161
	res.	51	81	7	8	.04	149
Males, 32 years or over	nonres.	51	81	7	9	.03	179
	res.	40	81	4	9	-.00	178
Females, 32 years or over	nonres.	54	81	6	8	.03	153
	res.	53	81	4	6	.02	143
Males, less than 12th grade	nonres.	48	81	3	4	.03	92
	res.	38	81	2	4	-.00	84
Females, less than 12th grade	nonres.	58	81	6	7	.07	81
	res.	65	81	6	6	.06	81
Males, 12th grade	nonres.	44	81	7	13	.02	130
	res.	41	81	3	9	.00	123
Females, 12th grade	nonres.	56	81	9	10	.06	136
	res.	51	81	5	9	.04	138
Males, more than 12th grade	nonres.	63	81	11	12	.07	106
	res.	46	81	2	7	.01	98
Females, more than 12th grade	nonres.	49	81	4	7	.03	97
	res.	40	81	3	8	-.00	83

positives than negatives among significant correlations (six out of seven for nonresidualized scores, four out of five for residualized scores) but not among all correlations (16 out of 27, 15 out of 27, respectively). In addition, they produce only trivial median correlations (r= 0.01, df=642, n.s.; r= 0.01, df=628 n.s., respectively). For both nonresidualized and residualized scores, demographic comparisons including sex differences yield only minor variations from this pattern of weak associations.

The subjective indicators of work involvement produce a somewhat higher proportion of positives among all correlations (61 out of 81 for nonresidualized scores, 51 out of 81 for residualized scores), likewise a higher proportion of positives among significant correlations (27 of 31, 13 of 18, respectively), and slightly higher median correlations (r=0.05, df=640, n.s.; r=0.03, df=593, n.s., respectively). For both nonresidualized and residualized scores, two measures of subjective reactions display the positive pattern most appreciably: time drags at work, and frequency of extra work. Again for both types of scores, demographic factors unearth few interesting subgroup differences, except that nonmarried males and males with a high school diploma do not conform at all to the positive trend.

Cell four examines the correlations in Tables 14 and 15 between eight measures of types of work activities and the three summary indices of degree of nonwork involvement. The correlations between the five indices of Quality of Employment and the three indices of nonwork involvement establish a consistent pattern, namely, positive and significant correlations for the

Table 14 (Cell 4)

Correlations between Measures of Types of Work Involvement and Degree of Nonwork Involvement for Total Sample

Types of Work Involvement	Nonwork Involvement: Memberships in selected organizations, non-res.	Memberships in selected organizations, res.	Overall nonwork activity, non-res.	Overall nonwork activity, res.	Overall nonwork involvement, non-res.	Overall nonwork involvement, res.
Quality of Employment Indices						
Overall Quality of Employment	.20** (646)	.04 (591)	.15** (631)	.00 (580)	.22** (626)	.06 (576)
Quality of Employment-Resources	.05 (634)	.00 (580)	-.02 (621)	-.08 (570)	.03 (616)	-.03 (566)
Quality of Employment-Financial rewards	.12** (640)	.02 (586)	.16** (626)	.06 (575)	.15** (621)	.08 (571)
Quality of Employment-Challenge	.20** (646)	.06 (591)	.17** (631)	.05 (580)	.23** (626)	.09* (576)
Quality of Employment-Comfort	.13** (645)	-.02 (591)	.02 (630)	-.09* (580)	.09* (625)	-.03 (576)
Measures of Effort						
Always a lot of work to do	.10** (645)	.03 (590)	.08* (630)	.01 (579)	.10* (625)	.02 (575)
Requires a lot of mental effort	.15** (646)	.06 (591)	.16** (631)	.07 (580)	.19** (626)	.09* (576)
Requires a lot of physical effort	-.10* (645)	-.02 (590)	-.17** (630)	-.11** (579)	-.18** (625)	-.10* (575)

Table 15 (Cell 4)

Correlation between Measures of Types of Work Involvement and Degree of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations		Significant Correlations		Median Correlations	
		No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation	N
		QUALITY OF EMPLOYMENT INDICES					
Total Sample	nonres.	14	15	11	11	.15**	631
	res.	10	15	1	2	.02	586
Males	nonres.	12	15	10	10	.15**	330
	res.	6	15	0	0	-.01	302
Females	nonres.	14	15	10	10	.15**	316
	res.	10	15	1	1	.06	278
Males, married	nonres.	13	15	9	9	.18**	256
	res.	8	15	0	0	.01	242
Females, married	nonres.	14	15	5	5	.14	170
	res.	11	15	1	1	.03	151
Males, not married	nonres.	10	15	0	0	.02	60
	res.	2	15	0	0	-.07	56
Females, not married	nonres.	14	15	8	8	.17*	139
	res.	9	15	0	0	.02	133
Males, under 32 years	nonres.	12	15	0	0	.09	145
	res.	1	15	0	0	-.07	136
Females, under 32 years	nonres.	15	15	10	10	.22**	160
	res.	11	15	3	3	.08	146
Males, 32 years or over	nonres.	12	15	7	7	.13	180
	res.	7	15	2	2	-.01	166
Females, 32 years or over	nonres.	14	15	1	1	.09	151
	res.	10	15	1	1	.02	140
Males, less than 12th grade	nonres.	11	15	5	5	.15	91
	res.	8	15	2	2	.03	80
Females, less than 12th grade	nonres.	14	15	4	4	.13	78
	res.	14	15	2	2	.12	69
Males, 12th grade	nonres.	14	15	3	3	.10	128
	res.	5	15	0	0	-.01	119
Females, 12th grade	nonres.	14	15	5	5	.16	139
	res.	10	15	0	0	.05	130
Males, more than 12th grade	nonres.	14	15	4	4	.06	101
	res.	1	15	0	1	-.10	95
Females, more than 12th grade	nonres.	12	15	2	2	.13	96
	res.	2	15	0	0	-.09	85

Table 15 (Cell 4 continued)

Correlation between Measures of Types of Work Involvement and Degree of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
MEASURES OF EFFORT							
Total	nonres.	6	9	6	9	.10*	625
Sample	res.	6	9	1	3	.02	575
Males	nonres.	6	9	4	6	.11	316
	res.	7	9	0	0	.01	298
Females	nonres.	6	9	2	5	.04	312
	res.	4	9	1	3	-.00	289
Males,	nonres.	6	9	3	6	.12	260
married	res.	4	9	0	0	-.00	241
Females,	nonres.	5	9	0	0	.02	171
married	res.	5	9	0	0	.01	155
Males,	nonres.	6	9	1	1	.08	61
not married	res.	7	9	0	0	.04	56
Females,	nonres.	6	9	2	5	.09	144
not married	res.	6	9	1	4	.02	133
Males, under	nonres.	4	9	0	0	-.01	150
32 years	res.	2	9	0	2	-.05	135
Females, under	nonres.	6	9	2	5	.12	160
32 years	res.	5	9	0	2	.06	149
Males, 32	nonres.	6	9	6	9	.17*	180
years or over	res.	6	9	2	2	.11	163
Females, 32	nonres.	4	9	1	3	-.01	152
years or over	res.	3	9	0	2	-.02	140
Males, less	nonres.	6	9	5	5	.25*	86
than 12th grade	res.	8	9	1	1	.13	81
Females, less	nonres.	3	9	0	0	-.05	79
than 12th grade	res.	4	9	0	0	-.02	70
Males,	nonres.	5	9	1	1	.01	126
12th grade	res.	3	9	0	0	-.01	119
Females,	nonres.	6	9	0	0	.04	135
12th grade	res.	6	9	0	0	.03	130
Males, more	nonres.	7	9	0	0	.11	103
than 12th grade	res.	4	9	0	0	-.04	97
Females, more	nonres.	6	9	1	1	.09	96
than 12th grade	res.	4	9	0	0	-.00	82

nonresidualized scores (median r=0.15, df=629, p < 0.01) but little relationship for the residualized scores (median r=0.02, df=584, n.s.). The pattern applies in particular to three indicators of Quality of Employment: overall, financial rewards, and challenge.

Based on nonresidualized scores, the positive relationships in different demographic subsamples between the indices of Quality of Employment and degree of nonwork involvement prove equally strong for males and females, but strongest among females 31 or under, and weakest among nonmarried males. The nonresidualized scores generate few median correlations of appreciable size, with one exception: a positive correlation among females with less than a high school diploma.

Three other measures of types of work activities tap the amount of energy that workers invest in their jobs and thus act as proxy measures for degree of involvement in work. They are: always a lot of work to do, requires a lot of mental effort, and requires a lot of physical effort. Yet these three items generate a conflicting pattern of findings when used to pit the positive against the negative hypothesis. The first two measures (always a lot of work to do, requires a lot of mental effort) produce positive correlations and thus support the positive hypothesis (nonresidualized scores only) whereas the third measure (requires a lot of physical effort) produces negative correlations and so favors the negative hypothesis (both types of scores).

Cell five concerns the correlations in Tables 16 and 17 between measures of types of work and types of nonwork activities. A low median correlation

Table 16 (Cell 5)

Correlations between Measures of Types of Work Involvement and Types of Nonwork Involvement for Total Sample

Types of Work Involvement	Nonwork Involvement: Sports & games		Cultural		Intellectual		Social		Family		Domesti
	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.
Overall Quality of Employment	.05 (648)	-.03 (632)	.16** (648)	.06 (597)	.08* (644)	-.04 (597)	.01 (636)	-.04 (604)	.03 (646)	.01 (645)	.13** (648)
Quality of Employment-Resources	-.05 (637)	-.08* (622)	.00 (637)	-.04 (587)	-.00 (633)	-.05 (587)	-.03 (626)	-.05 (594)	.01 (635)	-.01 (635)	.02 (637)
Quality of Employment-Financial Rewards	.13** (642)	.07 (628)	.17** (642)	.09* (592)	.18** (638)	.10* (592)	.06 (631)	.03 (598)	.00 (640)	-.02 (639)	.03 (642)
Quality of Employment-Challenge	.07 (648)	-.01 (632)	.20** (648)	.12** (597)	.06 (644)	-.05 (597)	.01 (636)	-.04 (604)	.06 (646)	.04 (645)	.15** (648)
Quality of Employment-Comfort	-.02 (647)	-.07 (631)	-.02 (647)	-.12** (597)	.03 (643)	-.07 (597)	.01 (635)	-.03 (604)	-.05 (645)	-.07 (644)	.07 (647)

Table 16 (Cell 5 continued)

Correlations between Measures of Types of Work Involvement and Types of Nonwork Involvement for Total Sample

	Nonwork Involvement				
	Political activities		# Close friends		# Neighbors know well enough to visit
Types of Work Involvement	non-res.	res.	non-res.	res.	nonresidual
Overall Quality of Employment	.19** (648)	.03 (609)	.03 (647)	.01 (643)	.04 (647)
Quality of Employment-Resources	.02 (637)	-.04 (599)	-.02 (636)	-.02 (632)	.03 (636)
Quality of Employment-Financial Rewards	.19** (642)	.11** (604)	-.01 (641)	-.02 (639)	.03 (641)
Quality of Employment-Challenge	.18** (648)	.04 (609)	.05 (647)	.02 (643)	.04 (647)
Quality of Employment-Comfort	.11** (647)	-.04 (609)	.02 (646)	.01 (642)	-.00 (646)

Table 17 (Cell 5)

Correlations between Measures of Types of Work Involvement and Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
QUALITY OF EMPLOYMENT INDICES							
Total Sample	nonres.	36	45	12	12	.03	642
	res.	22	45	4	6	-.01	632
Males	nonres.	37	45	12	12	..03	328
	res.	24	45	1	2	.000	328
Females	nonres.	32	45	10	10	..04	314
	res.	24	45	4	5	.01	287
Males, married	nonres.	34	45	11	11	.07	261
	res.	25	45	1	3	.02	267
Females, married	nonres.	31	45	7	7	.04	173
	res.	22	45	2	2	-.01	170
Males, not married	nonres.	21	45	0	1	-.02	61
	res.	14	45	0	3	-.08	60
Females, not married	nonres.	32	45	7	8	.04	142
	res.	22	45	2	4	-.01	130
Males, under 32 years	nonres.	23	45	5	5	.03	150
	res.	15	45	1	3	-.05	148
Females, under 32 years	nonres.	26	45	10	10	.03	164
	res.	17	45	4	5	-.02	151
Males, 32 years or over	nonres.	34	45	13	13	.08	179
	res.	26	45	3	6	.02	171
Females, 32 years or over	nonres.	32	45	3	3	.03	150
	res.	25	45	2	2	.01	136
Males, less than 12th grade	nonres.	28	45	9	10	.08	93
	res.	25	45	1	3	.03	85
Females, less than 12th grade	nonres.	25	45	3	4	.03	80
	res.	24	45	0	3	.03	70
Males, 12th grade	nonres.	28	45	4	4	.03	130
	res.	24	45	0	0	.01	129
Females, 12th grade	nonres.	31	45	7	7	.04	138
	res.	24	45	3	5	.00	125
Males, more than 12th grade	nonres.	35	45	0	0	.07	106
	res.	12	45	0	6	-.04	105
Females, more than 12th grade	nonres.	28	45	3	3	.05	96
	res.	18	45	0	4	-.04	97

notwithstanding (r=0.03, df=640, n.s.), the nonresidualized Quality of Employment indices tend to relate positively and significantly to the various types of nonwork activities: 36 positives out of 45 correlations; and 12 positives out of 12 significant correlations. These nonresidualized indices relate most positively and significantly to cultural activities and political activities; and display few associations with the other seven indices of nonwork involvement, hence the low median correlation. As for residualized scores, the Quality of Employment indicators appear generally unrelated to types of nonwork activities (22 positives out of 45 correlations; four positives out of six significant correlations; and a median correlation of -0.01 [df=630, n.s.]). With regard to demographic subgroups, the median correlations for nonresidualized scores differ little for men and women, and are greatest for married males, males over 31 and males with less than a high school diploma. In addition, nonmarried males produce the most negative median correlation based on residualized scores.

Tables 18 and 19 present a special set of correlations belonging to cell five. It includes measures of types of activities that provide the closest match possible within this dataset between work and nonwork. If the types of activities performed at work generalize to, or compensate for, off-the-job activities, or if nonwork has a similar effect on work, such trends should appear in Tables 18 and 19. The matching variables cover four areas: interpersonal involvement, supervision, mental effort, and physical effort.

The data in the tables represent a victory for the positive hypothesis. Among the 30 variable pairings based on nonresidualized scores, 23 produce positive correlations; and of the 10 significant correlations, all are positive. The median correlation is 0.06 (df=552, n.s.). The pattern

Table 18 (Cell 5)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Nonwork Involvement for Total Sample

Types of Work Involvement	Nonwork Involvement: Social non-res.	Social res.	Family non-res.	Family res.	# Close friends non-res.	# Close friends res.	# Neighbors know well enough to visit nonresidual	Youth group leader non-res.	Youth group leader res.	Intellectual non-res.	Intellectual res.
Interpersonal											
Not a one person job	.10* (635)	.09* (603)	.12** (644)	.13** (644)	.03 (645)	.03 (641)	.05 (645)				
Works in a group	.07 (637)	.06 (604)	.03 (646)	.03 (645)	-.06 (647)	-.07 (644)	-.05 (647)				
Number in workgroup	-.02 (546)	-.04 (518)	.03 (553)	.02 (552)	.14** (554)	.15** (551)	.07 (554)				
Workgroup cohesion	-.02 (547)	-.02 (519)	-.01 (553)	-.01 (553)	.01 (554)	.01 (551)	.06 (554)				
Work pace determined by others	.10* (635)	.06 (602)	.10* (645)	.10* (644)	.05 (646)	.04 (643)	.07 (646)				
Supervision											
Does respondent Supervise								.09* (647)	.04 (637)		
Number of workers R supervises								.02 (203)	-.00 (200)		
Mental Effort											
Requires a lot of mental effort										.11** (644)	.04 (597)
Requires learning new things										.07 (643)	.03 (595)
Requires creativity										.06 (644)	-.0[illegible] (597)

903

904

Table 18 (Cell 5 continued)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Involvement for Total Sample

Types of Work Involvement	Nonwork Involvement					
	Discussion group		Sports & games		Sports club or team	
	non-res.	res.	non-res.	res.	non-res.	res.
Mental effort						
Requires a lot of mental effort	.10* (647)	.06 (606)				
Requires learning new things	.12** (646)	.09* (604)				
Requires creativity	.10* (647)	.03 (606)				
Physical effort						
Requires a lot of physical effort			-.07 (647)	-.02 (631)	-.02 (646)	.04 (611)

Table 19 (Cell 5)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations		Significant Correlations		Median Correlations	
		No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation	N
				INTERPERSONAL			
Total Sample	nonres	15	20	5	5	.05	645
	res.	15	20	4	4	.05	645
Males	nonres.	14	20	6	6	.04	329
	res.	15	20	4	4	.03	328
Females	nonres.	17	20	3	3	.04	316
	res.	16	20	3	3	.04	282
Males, married	nonres.	16	20	6	6	.06	269
	res.	15	20	5	5	.06	269
Females, married	nonres.	14	20	4	4	.08	173
	res.	14	20	4	4	.07	172
Males, not married	nonres.	7	20	0	1	-.04	60
	res.	7	20	0	1	-.03	61
Females, not married	nonres.	13	20	1	1	.04	141
	res.	13	20	1	1	.03	134
Males, under 32 years	nonres.	13	20	2	3	.03	150
	res.	11	20	2	3	.03	138
Females, under 32 years	nonres.	12	20	0	0	.03	178
	res.	12	20	0	0	.01	154
Males, 32 years or over	nonres.	16	20	4	4	.05	149
	res.	14	20	3	3	.04	148
Females, 32 years or over	nonres.	17	20	2	2	.07	136
	res.	17	20	2	2	.06	129
Males, less than 12th grade	nonres.	15	20	2	2	.06	93
	res.	13	20	2	2	.07	93
Females, less than 12th grade	nonres.	13	20	2	2	.05	73
	res.	14	20	2	2	.06	75
Males, 12th grade	nonres.	12	20	1	1	.07	130
	res.	12	20	0	0	.07	105
Females, 12th grade	nonres.	17	20	2	2	.06	121
	res.	16	20	2	2	.06	139
Males, more than 12th grade	nonres.	14	20	1	1	.02	103
	res.	12	20	1	1	.02	99
Females, more than 12th grade	nonres.	12	20	1	1	.01	97
	res.	10	20	1	1	.01	88

Table 19 (Cell 5 continued)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Nonwork Involvement for Total Sample and Subsamples

		Positive Correlations		Significant Correlations		Correlation Coefficients			
Samples	Types of Scores	No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation[a]	N	Correlation[b]	N
				SUPERVISION					
Total	nonres.	2	2	1	1	.09*	647	.02	201
Sample	res.	1	2	0	0	.04	637	−.00	200
Males	nonres.	1	2	1	1	.10*	330	−.01	126
	res.	1	2	0	0	.06	324	−.02	124
Females	nonres	0	2	0	0	−.00	317	−.03	77
	res.	0	2	0	0	−.04	313	−.04	76
Males, married	nonres.	2	2	0	0	.09	269	.03	110
	res.	2	2	0	0	.05	264	.03	109
Females, married	nonres.	1	2	0	0	−.02	173	.03	44
	res.	1	2	0	0	−.05	170	.02	44
Males, not married	nonres.	1	2	0	0	.16	61	−.15	16
	res.	1	2	0	0	.14	60	−.20	15
Females, not married	nonres.	1	2	0	0	.01	143	−.12	33
	res.	0	2	0	0	−.04	142	−.12	32
Males, under 32 years	nonres.	2	2	0	0	.09	150	.06	35
	res.	2	2	0	0	.04	149	.03	34
Females, under 32 years	nonres.	0	2	0	0	−.01	164	−.10	25
	res.	0	2	0	0	−.04	161	−.18	25
Males, 32 years or over	nonres.	1	2	0	0	.06	180	−.04	91
	res.	1	2	0	0	.05	175	−.04	90
Females, 32 years or over	nonres.	2	2	0	0	.00	153	.01	52
	res.	1	2	0	0	−.05	152	.04	51
Males, less than 12th grade	nonres.	2	2	0	0	.18	92	.01	20
	res.	1	2	0	0	.16	90	−.05	20
Females, less than 12th grade	nonres.	c	c	c	c	c		c	
	res.	0	2	0	0	−.03	78	−.12	14
Males, 12th grade	nonres.	2	2	0	0	.01	130	.04	49
	res.	1	2	0	0	−.01	129	.03	49
Females, 12th grade	nonres.	0	2	0	0	−.03	139	−.02	34
	res.	0	2	0	0	−.06	138	−.03	34
Males, more than 12th grade	nonres.	1	2	0	0	.11	106	−.02	56
	res.	1	2	0	0	.06	104	−.03	55
Females, more than 12th grade	nonres.	1	2	0	0	.02	97	−.09	27
	res.	0	2	0	0	−.06	95	−.13	27

a This column contains correlations between the item "Youth Group Leader" (nonresidualized and residualized) and "Does respondent Supervise"

b This column contains correlations between the item "Youth Group Leader" (nonresidualized and residualized) and "Number of Workers respondent Supervises"

c No correlation reported since scores do not vary on the item "Youth Group Leader."

Table 19 (Cell 5 continued)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
				MENTAL EFFORT			
Total Sample	nonres.	6	6	4	4	.10*	647
	res.	5	6	1	1	.03	595
Males	nonres.	6	6	3	3	.09	328
	res.	5	6	0	0	.05	308
Females	nonres.	5	6	3	3	.07	315
	res.	4	6	1	1	.04	289
Males, married	nonres.	6	6	3	3	.11	269
	res.	6	6	0	0	.05	252
Females, married	nonres.	3	6	1	1	-.02	173
	res.	3	6	0	0	-.03	158
Males, not married	nonres.	3	6	0	0	-.00	61
	res.	2	6	0	0	-.05	57
Females, not married	nonres.	5	6	0	0	.12	144
	res.	0	6	0	0	.08	135
Males, under 32 years	nonres.	4	6	0	0	.05	150
	res.	3	6	0	0	-.02	139
Females, under 32 years	nonres.	5	6	2	2	.08	163
	res.	4	6	1	1	.05	151
Males, 32 years or over	nonres.	6	6	4	4	.18*	178
	res.	6	6	2	2	.10	171
Females, 32 years or over	nonres.	4	6	0	0	.05	153
	res.	4	6	0	0	.02	143
Males, less than 12th grade	nonres.	6	6	2	2	.18	92
	res.	6	6	1	1	.15	85
Females, less than 12th grade	nonres.	3	6	0	0	-.09	80
	res.	3	6	0	0	-.04	71
Males, 12th grade	nonres.	4	6	1	1	.05	129
	res.	4	6	0	0	.00	121
Females, 12th grade	nonres.	4	6	0	0	.05	137
	res.	4	6	0	0	.01	133
Males, more than 12th grade	nonres.	4	6	0	0	.04	106
	res.	0	6	0	0	-.09	101
Females, more than 12th grade	nonres.	6	6	0	0	.11	97
	res.	5	6	0	0	.04	84

Table 19 (Cell 5 continued)

Correlations between Selected Measures of Types of Work Involvement and Corresponding Measures of Types of Nonwork Involvement for Total Sample and Subsamples

		Positive Correlations		Significant Correlations		Correlation Coefficients			
Samples	Types of Scores	No. of positive correlations	No. of correlations	No. of significant positive correlations	No. of significant correlations	Correlation[a]	N	Correlation[b]	N
				PHYSICAL EFFORT					
Total Sample	nonres.	0	2	0	0	-.07	647	-.02	646
	res.	1	2	0	0	-.02	631	.04	611
Males	nonres.	0	2	0	0	-.05	330	-.02	329
	res.	2	2	0	0	.05	324	.07	311
Females	nonres.	0	2	0	1	-.12*	317	-.04	317
	res.	0	2	0	1	-.12*	307	-.02	300
Males, married	nonres.	0	2	0	0	-.03	269	-.02	268
	res.	2	2	0	0	.07	265	.08	254
Females, married	nonres.	1	2	0	0	-.11	173	.06	172
	res.	1	2	0	0	-.10	168	.08	163
Males, not married	nonres.	1	2	0	0	-.15	61	.01	61
	res.	1	2	0	0	-.10	59	.02	57
Females, not married	nonres.	0	2	0	0	-.14	143	-.16	144
	res.	0	2	0	0	-.14	138	-.15	136
Males, under 32 years	nonres.	1	2	0	0	-.02	151	.02	150
	res.	2	2	0	0	.05	147	.07	139
Females, under 32 years	nonres.	0	2	0	0	-.11	165	-.05	164
	res.	0	2	0	0	-.11	161	-.02	156
Males, 32 years or over	nonres.	0	2	0	0	-.09	179	-.04	179
	res.	2	2	0	0	.02	177	.07	172
Females, 32 years or over	nonres.	0	2	0	0	-.13	152	-.02	153
	res.	0	2	0	0	-.12	146	-.02	144
Males, less than 12th grade	nonres.	1	2	0	0	.07	92	-.03	91
	res.	2	2	0	0	.16	90	.08	85
Females, less than 12th grade	nonres.	2	2	0	0	.12	80	.02	79
	res.	2	2	0	0	.13	76	.01	73
Males, 12th grade	nonres.	0	2	0	0	-.05	130	-.03	130
	res.	1	2	0	0	-.01	128	.02	124
Females, 12th grade	nonres.	0	2	0	0	-.13	139	-.09	139
	res.	0	2	0	0	-.12	139	-.10	136
Males, more than 12th grade	nonres.	2	2	0	0	.09	106	.13	106
	res.	2	2	1	1	.19	105	.20*	101
Females, more than 12th grade	nonres.	2	2	0	0	.06	97	.09	97
	res.	2	2	0	0	.06	91	.15	89

[a] This column contains correlations between the item "Job Requires a Lot of Physical Effort" and "Sports & Games" (nonresidualized and residualized)

[b] This column contains correlations between the item "Job Requires a Lot of Physical Effort" and "Sports Club or Team" (nonresidualized and residualized)

created by the residualized scores is similar: 22 positives out of 30 correlations, five positives out of five significant correlations, and a median correlation of 0.04 (df=595, n.s.). Because residualization has little effect on the findings concerning matched activities, residualized scores are not further reported for cell five. Two of the four areas (interpersonal involvement, mental effort) contain enough correlations to permit comparisons among areas as to their degree of support for the positive hypothesis. The correlations in these two areas indicate somewhat more support for the hypothesis from measures of mental effort than from those of interpersonal involvement. The specifics are as follows: for mental effort (six positives out of six correlations, four positives out of four significant correlations, and a median correlation of 0.10 [df=645, $p < 0.05$]); and for interpersonal involvement (15 positives out of 20 correlations, five positives out of five significant correlations, and a median correlation of 0.05 [df=643, n.s.]). With regard to the other two areas, the measures of supervision offer some support for the positive hypothesis but the measures of physical effort offer none.

Demographic factors have no great impact on these findings on matched activities. Sex differences are negligible for the data on all four areas combined, although nonmarried males display the pattern of positive matching somewhat more weakly. With the areas considered separately, the data on mental effort support the positive hypothesis somewhat more among men over 31 and men with less than a high school diploma. The data on interpersonal involvement, by contrast, exhibit less support among nonmarried males than any of the other subgroups. Otherwise, the small numbers of correlations and cases available warrant few comments on demographic differences.

Cell seven tests the positive and negative hypotheses in terms of correlations between nine measures of reactions to work involvement and three measures of degree of nonwork involvement. According to Tables 20 and 21 the nonresidualized scores support the positive hypothesis since they generate an overwhelmingly positive pattern of both correlations (25 out of 27) and significant correlations (10 out of 10). The median correlation of 0.07, nontheless, falls short of statistical significance (df=623, n.s.). Although the three measures of nonwork involvement produce quite similar results, two of the measures of reactions to work involvement differ from the remaining six: absence of job depression correlates very strongly and positively with nonwork involvement; and satisfaction with resources correlates weakly and negatively with nonwork involvement.

The residualized scores provide weak support for the negative hypothesis. The correlations follow a negative pattern (seven positives out of 27 correlations), as do the significant correlations (zero positives out of two). In addition, the median correlation proves quite trivial (r=-0.02, df=572, n.s.). The negative pattern appears somewhat stronger for the index of overall nonwork activity than for the other two measures of nonwork involvement; and also somewhat stronger for satisfaction with challenge and with resources than for absence of job depression.

As regards demographic subgroups, the nonresidualized scores produce a larger (positive) median correlation among females than males and largest (positive) median correlations among females 31 and under,

Table 20 (Cell 7)

Correlations between Measures of Personal Reactions to Work and Nonwork for Total Sample

Work Reaction	Nonwork Involvement					
	Memberships in selected organizations		Overall nonwork activity		Overall nonwork involvement	
	non-res.	res.	non-res.	res.	non-res.	res.
Facet-free job satisfaction	.10* (646)	.01 (590)	.01 (632)	-.09* (580)	.05 (627)	-.06 (576)
Job satisfaction w/coworkers	.10* (645)	.04 (589)	.05 (630)	-.01 (578)	.10* (625)	.03 (574)
Job satisfaction w/challenge	.12** (644)	-.01 (588)	.03 (630)	-.08** (578)	.08 (625)	-.05 (574)
Job satisfaction w/comfort	.09** (644)	-.01 (588)	.00 (630)	-.07 (578)	.08 (625)	-.01 (574)
Job satisfaction w/resources	.02 (645)	-.05 (589)	-.02 (630)	-.08 (578)	.02 (625)	-.05 (574)
Job satisfaction w/financial reward	.05 (644)	-.01 (588)	.02 (630)	-.08 (578)	.07 (625)	-.02 (574)
Facet-specific job satisfaction	.10** (644)	-.01 (588)	.02 (630)	-.08 (578)	.09* (625)	-.03 (574)
Absence of job depression	.15** (645)	.06 (589)	.16** (631)	.08 (579)	.18** (626)	.08 (575)
Financial equity	.08 (602)	.04 (552)	-.02 (592)	-.07 (545)	.00 (587)	-.06 (541)

Table 21 (Cell 7)

Correlations between Measures of Personal Reactions to Work and Nonwork for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlation	Median Correlations: N
Total Sample	nonres.	25	27	10	10	.07	625
	res.	7	27	0	2	-.02	574
Males	nonres.	22	27	3	3	.05	330
	res.	4	27	0	5	-.07	298
Females	nonres.	24	27	10	10	.08	307
	res.	17	27	1	1	.02	280
Males, married	nonres.	24	27	2	2	.07	258
	res.	4	27	0	3	-.06	231
Females, married	nonres.	22	27	7	7	.07	167
	res.	22	27	1	1	.03	153
Males, not married	nonres.	8	27	0	0	-.04	61
	res.	7	27	0	1	-.10	53
Females not married	nonres.	21	27	1	1	.07	143
	res.	6	27	0	1	-.05	131
Males, under 32 years	nonres.	24	27	4	4	.09	148
	res.	7	27	1	2	-.05	136
Females, under 32 years	nonres.	24	27	8	8	.12	159
	res.	14	27	0	0	.03	145
Males, 32 years or over	nonres.	9	27	0	0	-.02	172
	res.	1	27	0	4	-.09	166
Females, 32 years or over	nonres.	23	27	4	4	.06	153
	res.	18	27	2	2	.04	135
Males, less than 12th grade	nonres.	16	27	0	0	.02	86
	res.	2	27	0	5	-.12	82
Females, less than 12th grade	nonres.	18	27	4	4	.07	80
	res.	21	27	1	2	.05	69
Males, 12th grade	nonres.	21	27	2	2	.06	130
	res.	9	27	0	0	-.02	119
Females 12th grade	nonres.	26	27	10	10	.13	125
	res.	20	27	2	2	.07	129
Males, more than 12th grade	nonres.	20	27	4	4	.09	103
	res.	6	27	0	0	-.07	97
Females, more than 12th grade	nonres.	21	27	1	1	.06	91
	res.	4	27	0	0	-.09	85

and females with a high school diploma. The residualized scores generate a slightly larger (negative) median correlation among males than females, and largest (negative) median correlations among unmarried males, males over 31, males with less than a high school diploma, and females with more than a high school diploma.

The correlations between eight measures of personal reactions to work involvement and nine measures of types of nonwork involvement belong to cell eight and appear in Tables 22 and 23. The nonresidualized scores display a somewhat positive pattern: mainly positives among all correlations (46 out of 81) and among all significant correlations (19 out of 23), but a wholly trivial median correlation (r=0.02, df=601, n.s.). The positive pattern seems sharpest for three types of nonwork involvement (cultural activities, domestic activities, and political activities) and least apparent for three others (social activities, intellectual activities, and sports & games). As regards measures of reactions to work, absence of job depression adheres to the positive trend much more strongly than any of the satisfaction measures, of which two (satisfaction with resources, satisfaction with financial rewards) do not adhere to it at all.

The residualized scores, however, tell a somewhat different story. They conform to a negative pattern to roughly the degree to which the nonresidualized scores exhibit a positive trend: mostly negative correlations (35 positives out of 81), only two positives among the 20 significant correlations, and again a trivial median correlation (r=-0.01, df=593, n.s.). Four types of nonwork activities conform to the negative trend (social activities, intellectual activities, family activities, and sports & games),

Table 22 (Cell 8)

Correlations between Measures of Personal Reaction to Work and Types of Off-the-job Involvement for Total Sample

Work Reaction	Nonwork Involvement											
	Sports & games		Cultural		Intellectual		Social		Family		Domestic	
	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.	non-res.	res.
Facet-free job satisfaction	-.04 (649)	-.09* (633)	.05 (649)	-.01 (597)	-.06 (645)	-.14** (597)	-.07 (637)	-.11** (604)	.01 (647)	-.00 (646)	.09* (649)	.03 (643)
Job satisfaction w/coworkers	.05 (647)	.03 (631)	.03 (647)	.00 (595)	-.00 (643)	-.05 (595)	-.00 (635)	-.04 (602)	-.00 (645)	-.01 (644)	.01 (647)	-.01 (641)
Job satisfaction w/challenge	-.03 (647)	-.09* (631)	.11** (647)	.05 (595)	-.04 (643)	-.13** (595)	-.09* (635)	-.13** (602)	-.03 (645)	-.04 (644)	.11** (647)	.04 (641)
Job satisfaction w/comfort	-.05 (647)	-.09* (631)	.04 (647)	-.01 (595)	-.04 (643)	-.10* (595)	-.04 (635)	-.08 (602)	-.06 (645)	-.08* (644)	.08* (647)	.04 (641)
Job satisfaction w/resources	-.10* (647)	-.12** (631)	.03 (647)	.01 (595)	-.00 (643)	-.04 (595)	-.04 (635)	-.07 (602)	-.04 (645)	-.06 (644)	.02 (647)	-.01 (641)
Job satisfaction w/financial reward	-.01 (647)	-.05 (631)	.08* (647)	.03 (595)	-.04 (643)	-.11** (595)	-.08* (635)	-.11** (602)	-.03 (645)	-.03 (644)	.07 (647)	.02 (641)
Facet-specific job satisfaction	-.04 (647)	-.09* (631)	.09* (647)	.03 (595)	-.04 (643)	-.12** (595)	-.07 (635)	-.11** (602)	-.04 (645)	-.06 (644)	.09* (647)	.03 (641)
Absence of job depression	.09* (648)	.02 (632)	.08* (648)	.02 (596)	.08* (644)	.04 (596)	.01 (636)	-.03 (603)	.08* (646)	.06 (645)	.16** (648)	.11** (642)
Financial Equity	.02 (604)	-.02 (592)	-.01 (604)	-.04 (558)	-.08 (601)	-.11* (560)	-.10* (596)	-.11** (566)	-.00 (603)	-.01 (603)	.10* (604)	.05 (601)

Table 22 (Cell 8 continued)

Correlations between Measures of Personal Reaction to Work and Types of Nonwork Involvement for Total Sample

	Nonwork Involvement				
	Political activities		# Close friends		# Neighbors know well enough to visit
Work Reaction	non-res.	res.	non-res.	res.	nonresidual
Facet-free job satisfaction	.05 (649)	-.04 (609)	.04 (648)	.03 (644)	-.03 (648)
Job satisfaction w/coworkers	.02 (647)	-.04 (609)	.03 (646)	.02 (642)	.05 (646)
Job satisfaction w/challenge	.12** (647)	.00 (609)	.02 (646)	.01 (642)	.01 (646)
Job satisfaction w/comfort	.13** (647)	.03 (609)	-.02 (646)	-.02 (642)	.04 (646)
Job satisfaction w/resources	.04 (647)	-.03 (609)	-.04 (646)	-.03 (642)	.01 (646)
Job satisfaction w/financial reward	.08* (647)	.01 (609)	.04 (646)	.02 (642)	-.01 (646)
Facet-specific job satisfaction	.12** (647)	.00 (609)	.00 (646)	.00 (642)	.03 (646)
Absence of job depression	.10* (648)	.01 (608)	.11** (647)	.09* (643)	.01 (647)
Financial equity	-.05 (604)	-.08* (570)	.02 (603)	-.01 (601)	.04 (603)

Table 23 (Cell 8)

Correlations between Measures of Personal Reaction to Work and Types of Nonwork Involvement for Total Sample and Subsamples

Samples	Types of Scores	Positive Correlations: No. of positive correlations	Positive Correlations: No. of correlations	Significant Correlations: No. of significant positive correlations	Significant Correlations: No. of significant correlations	Median Correlations: Correlations	Median Correlations: N
Total Sample	nonres.	46	81	19	23	.02	603
	res.	35	81	2	20	-.01	595
Males	nonres.	49	81	6	10	.01	330
	res.	22	81	0	14	-.03	307
Females	nonres.	49	81	14	20	.03	316
	res.	37	81	5	13	-.00	306
Males, married	nonres.	52	81	5	8	.03	262
	res.	32	81	0	9	-.03	254
Females, married	nonres.	53	81	3	6	.03	173
	res.	40	81	1	4	-.00	162
Males, not married	nonres.	22	81	0	6	-.09	61
	res.	16	81	0	9	-.12	61
Females, not married	nonres.	44	81	4	6	.01	142
	res.	39	81	1	10	-.01	141
Males, under 32 years	nonres.	43	81	7	9	.01	151
	res.	24	81	0	4	-.06	137
Females, under 32 years	nonres.	44	81	12	13	.01	164
	res.	34	81	5	10	-.02	155
Males, 32 years or over	nonres.	41	81	0	5	.00	180
	res.	24	81	0	9	-.04	168
Females, 32 years or over	nonres.	48	81	1	2	.03	152
	res.	44	81	0	1	.01	142
Males, less than 12th grade	nonres.	42	81	0	10	.01	93
	res.	29	81	0	14	-.05	85
Females, less than 12th grade	nonres.	39	81	0	5	-.01	81
	res.	37	81	1	6	-.01	70
Males, 12th grade	nonres.	46	81	4	5	.02	130
	res.	37	81	1	5	-.01	123
Females, 12th grade	nonres.	59	81	11	14	.06	138
	res.	50	81	2	7	.03	138
Males, more than 12th grade	nonres.	61	81	5	6	.05	106
	res.	30	81	1	6	-.03	100
Females, more than 12th grade	nonres.	54	81	5	7	.03	97
	res.	40	81	2	8	-.00	83

whereas the remainder exhibit a positive trend. As for personal reactions to work, five measures sustain the negative pattern most clearly (facet-free job satisfaction, facet-specific job satisfaction, satisfaction with challenge, satisfaction with comfort, and financial equity), and only absence of job depression displays a positive trend.

The small sizes of the correlations in cell eight justify few comments about variations among demographic subgroups, aside from the observation that nonmarried males generate the largest median negative correlations for both nonresidualized and residualized scores.

Discussion

Univariate Distributions for Measures of Types of Nonwork Activities

Activities: The present study found that among workers, the most frequent off-the-job activities included talking on the phone with friends, getting in touch with relatives and with friends, fixing things around the home, reading a magazine article, going out with the family, and gardening or working around the yard. The domestic theme that pervades such rankings also appeared in an earlier U.S. study. Converse and Robinson (1972) presented data on time-budgets from a national cross-section of Americans in 1965-66. They found that top rankings for time spent on leisure activities belonged to watching television, visiting, and reading for three groups of employed men (executive/professional, white-collar, and blue-collar); and to visiting, watching television, and reading for two groups of employed women (white-collar, blue-collar).

Several overseas studies have likewise accorded primacy to domestically-oriented leisure. Matthews and Abu-Laban (1959) noted that more of the foreign teachers in their study in Kuwait reported reading, visiting with friends, and listening to radio programs and music than engaging in any other leisure activity. Goldthorpe et al. (1969) observed that the husbands in their English sample of affluent blue-collar workers spent more spare time on two activities than on any others: chores and odd jobs including gardening; and home-based leisure, most notably watching television. The patterns of sociability of these husbands, moreover, commonly involved kin and neighbors. A comparison sample of white-collar working husbands revealed an equally strong trend toward family-centered leisure. Willmott (1971) found that, among an English sample of married male employees aged 30 and over, the most frequent and the most preferred nonwork activities were home-based: home decorations, gardening, listening to music, car maintenance, and reading. Of every 10 activities outside the home, six involved other family members. Brown, Brannen, Cousins and Samphier (1973) observed that for a sample of English shipbuilding workers normal weekly leisure activities were largely home-based (watching television, reported by 78%; doing jobs in the house, 64%, reading, 45%; gardening, 42%; activities with children, 30%; and hobbies, 21%) and family-based (visiting relatives, 45%; shopping, 39%; and family outings, 27%).

As for distributional data on number of close friends, 41% of the respondents in the present study reported fewer than five close friends. In an English study of just two occupational groups, arcitects and railway men, Salaman (1974) obtained the somewhat lower figures of 2% and 33%, respectively, on a parallel question about number of best friends.

Organizations. Whereas the present study measured organizational involvement in terms of number of organizations belonged to, certain other investigators have explored more elaborate measures. Goldthorpe et al. (1969), for example, ascertained whether their respondents were organizational committee members or office holders, as well as members. Hagedorn and Labovitz (1968) included questions about membership, financial support, active participation, office holding, and committee membership, in their study of participation. Wilensky (1961) recommended that membership and meeting attendance be supplemented by a measure of hours per month spent on all organizational activity including meetings, phone calls, committee work, and money-raising (cf. Meissner, 1970). Such methodological refinements suggest that the number of organizations belonged to may tap variety as much as amount of organizational involvement (cf. Matthews and Abu-Laban, 1959; Kelly, 1972).

The available literature on the (simple) measure of number of organizational memberships reinforces the impression conveyed by Table 7 that most workers belong to few organizations: 33% belonged to no organizations, 29% to one organization, and 38% to two or more. The univariate distribution in Table 7, it should be stressed, discounts membership in labor unions, professional associations, parent-teacher associations, youth groups and other organizations not included in the list shown in Table 6. By way of comparison, Goldthorpe et al. (1969) reported that the English working husbands in their sample belonged to formal associations (excluding trade unions and general work clubs) as follows: 33% of the white-collar husbands belonged to no association, versus 52% for manual husbands; 28% of the white-collar husbands belonged to one association, versus 33% for manual husbands; and 39% of the white-collar husbands belonged to two or more associations, versus 14% for manual husbands. Brown et al. (1973) noted that among their sample of English shipbuilders 32% belonged to no clubs or organizations (excluding trade unions), 52% belonged to one, and 16% to two or more.

As regards the types of organizations with which workers affiliate most frequently, the discussion must proceed at the level of individual organizations since empirically validated taxonomies of organizations do not abound. The literature, in short, has not provided the type of meaningful taxonomy sought unsuccessfully from the data reduction procedures in the present study. Although Wilensky (1961) has suggested a number of potential taxonomic dimensions for organizations (size, control structure, membership base, scope of activities, purposes or functions, auspices, degree of internal solidarity, degree of conflict with environment, etc.), he has acknowledged the paucity of studies that systematically compare organizations on these and similar dimensions. The present study found more than twice as many workers belonged to a church or synogogue (47%) as to any other organization (e.g., sports club or team, 21%; labor union, 19%). By comparison, Meissner (1970) found that 46% of workers in a large manufacturing firm were members of a church, but Brown et al. (1973) recorded that only 5% of the English shipbuilders in their sample belonged to a church group.

Measures of Resources for Nonwork Involvement

All measures of the resources needed for nonwork involvement were required to meet two tests. They first had to correlate positively with other measures of the same resource, a requirement that posed no problem for measures of three resources--energy, health, and money--but which eliminated several measures of time as a resource: number of hours worked per week, time taken to travel to work, number of paid vacation days, and number of problems with transportation to work.

All measures, secondly, had to correlate positively with at least one of the 12 nonresidualized measures of nonwork involvement. This second criterion eliminated one measure of energy as a resource (nonobesity) and two measures of income as a resource (satisfaction with fringe benefits and satisfaction with pay). Though not necessary for the analysis, correlations were also computed between the four measures of time as a resource that failed at test one and the 12 measures of nonwork involvement. Few significantly positive correlations emerged: two for number of paid vacation days (intellectual activities, memberships in organizations); one for number of hours worked per week (social activities) and for time taken to travel to work (sports & games); and zero for number of problems with transportation to work. The paucity of significantly positive correlations between the measures of time as a resource that failed test one and the measures of nonwork involvement demonstrates that the initial test did not eliminate any strong measures of resources.

Clearly the data do not establish a connection between all plausible measures of resources and nonwork involvement. Many likely measures of resources failed to meet the two statistical criteria. By stipulation, all of the finally selected measures of resources correlated positively and significantly with at least one measure of nonwork involvement. What is surprising is the absence from the list of a number of other plausible measures of resources that failed one or other of the tests (e.g., number of hours worked per week).

Despite the availability in the interview of a wide variety of potential measures of resources, at least one major resource did not receive adequate measurement, namely, social skills. Had measures of social skills been available and thus included in the multiple regressions, some of the correlations between social dimensions of work and nonwork activities based on residualized scores may have been weakened. Nor was there a direct measure of another resource, occupationally-based social and organizational connections. In addition, certain specific omissions from the measures of resources should be recognized. Measures of amount and adequacy of family income, for example, would possibly have bolstered the measurement of money as a resource.

Summary of Cells One to Nine

The data from the cells depicted in Table 1 (Part I) may be viewed in the context of other relevant studies (described in Part I) in order to evaluate the positive and negative hypotheses. For a variety of reasons, the pattern of evidence from present and past sources is neither comprehensive nor entirely consistent. The present study, on the one hand, contains no measures of subjective reactions to nonwork, hence the absence of any data for cells three, six and nine. On the other hand, the available literature contains studies that differ widely with regard to historical period, geographic location (including international as well as intranational comparisons), population sampled, sampling technique, types of measures, statistical analyses, use of significance tests, etc. Studies in the literature, moreover, have tended to focus on either the relationship between work and leisure or that between work and family, but not on the two relationships jointly. Such imperfections notwithstanding, the hope remains that some of the limitations of the present and previous studies will cancel each other out, thereby allowing a more coherent picture to emerge of the relationship between work and nonwork.

The sequential review of data for cells one, two, four, five, seven and eight concentrates initially on nonresidualized measures of nonwork activities as applied to the whole sample. Considered next are the variations in correlational patterns for different demographic subsamples. Discussion of possible causal inferences, the third item on the agenda, includes but is not limited to consideration of the residualized scores for nonwork. The discussion ends with a brief statement of major conclusions.

Choices among measures of degree of involvement do affect the results in cell one. Two types of measures tap degree of involvement in work: objective (time devoted to work) and subjective (feeling involved, investing effort, etc.). The three indices of degree of involvement in nonwork are all objective, though in a different sense: they measure overall levels of activity summed across various categories of nonwork.

The data in cell one display no relationship between degree of work involvement (measured objectively) and degree of nonwork involvement (also measured objectively); but they do show a positive relationship between degree of work involvement (measured subjectively) and degree of nonwork involvement (measured objectively).

As regards relevant studies in the literature, Clark and his colleagues (Clark & Gecas, 1977; Clark et al., 1977) found marginally negative relationships among husbands between degree of involvement in work (measured objectively in terms of number of hours) and degree of involvement in family roles (also measured objectively in terms of hours).[1] Grubb (1975) obtained no compelling relationship between degree of involvement in work (measured subjectively) and degree of involvement in nonwork (measured objectively in terms of participation in nine favorite leisure activities).

Cell two concerns the relationship between degree of involvement in work and participation in types of nonwork activities. The correlational pattern obtained suggests no relationship for objective measures of degree of work involvement but a positive relationship for subjective measures. The similarity between the data obtained for cells one and two in this study (the trends being weaker in cell two) should cause no surprise since the measures of degree of involvement in nonwork used in cell one are a summated version of the measures of types of nonwork activities used in cell two.

---

[1] The most precise comparison between Clark's findings and those in the present study concerns the correlation in this study between degree of involvement in work (measured objectively in terms of number of hours on main job) and degree of involvement in family activities (measured objectively in terms of recency of two basic family activities). Based on the subsample of husbands, the correlation (which is not shown in any of the tables) proves nonsignificant (r=0.07, df=267, n.s.) in contrast to Clark's marginally negative relationships.

The findings for cell two make little contact with those in the literature. Though not tested in cell two in the present study, there exists one version of the positive hypothesis that is conceptually associated with cell two. It asserts that employees highly involved in their work participate in nonwork activities that are similar to their activities at work, whereas employees less involved in their work do not exhibit this pattern.[2] As noted earlier (Part 1), a study by Parker (1965) supported this version of the positive hypothesis.

Despite the fact that the present study contributes no new findings to cell three, it bears repetition that earlier studies have found few consistent relationships between degree of involvement in work and subjective reactions to nonwork, although the relationships do seem more frequently to be negative when subjective as opposed to objective measures of degree of work involvement are used.

The positive and negative hypotheses in cell four concern whether the amount of effort that jobs demand of workers is related to their degree of involvement in nonwork. The correlations prove negative for physical effort, positive for mental effort, and positive for a third but less specific measure, always a lot of work to do. Two studies in the literature have examined the relationship between physically fatiguing jobs and degree of involvement in leisure. Larrue's (1965) study detected no association between the two concepts; yet, Lundahl (1971) reported a negative relationship, especially for cultural and intellectual activities, which are also two of the three types of leisure activities that correlate most negatively with physical effort in the present study. (The latter correlations do not appear in any of the tables).

---

[2] The best test in the present study of this variant of the positive hypothesis actually occurs in cell five. With level of education acting as a proxy for degree of involvement in work, the correlations in cell five between similar activities undertaken on and off the job do not appear to be consistently or significantly greater at higher levels of education (Table 19).

Cell five concerns the possible match between activities at work and those away from work. The correlations in the cell support the positive hypothesis overall and also for three of the four subcategories of activities: mental effort, interpersonal involvement, and supervision, but not physical effort. By comparison, Meissner's (1971) study supported the positive hypothesis for interpersonal involvement, although Hagedorn and Labovitz' (1968) data favored the negative hypothesis. In addition, Meissner's investigation obtained positive findings for degree of discretion or autonomy (a fifth subcategory) as did Hagedorn and Labovitz' for supervision.

Cell six holds little interest here. It does not concern the positive and negative hypotheses. It cannot be investigated with the present dataset, and it contains no studies from the pertinent literature.

Cell seven involves correlations between subjective reactions to work and degree of involvement in nonwork. The present data support the positive hypothesis quite consistently, whereas the studies in the literature suggest no relationship in three instances (Grubb, 1975; Larrue, 1965; and Seeman, 1967) and a negative relationship in a fourth (Matthews & Abu-Laban, 1959). In particular, since Seeman's study concerned involvement in work and extent of political knowledge, its failure to find a relationship may be compared to the significantly positive correlations in the present study between subjective reactions to work and participation in political activities. (Strictly, these latter correlations belong to cell eight and thus appear in Table 22).

Cell eight does not concern the positive or negative hypotheses but the correlations between subjective reactions to work and types of nonwork activities remain of considerable interest. The empirical parallel between the results for cells seven and eight resembles the corresponding parallels noted for cells one and two. In each case, the measures of degree of involvement in nonwork (cells one and seven) are a summated version of the measures of types of nonwork activities (cells two and eight). According to the present data for cell eight, favorable reactions to work tend to be positively associated with participation in various nonwork activities. This is especially true of cultural, domestic, and political activities, but not at all true of sports & games.

As regards relevant literature, four studies (Bacon, 1975; Hanhart, 1964; Larrue, 1965; and Matthews & Abu-Laban, 1959) for the most part detected no association between work satisfaction and various types of nonwork activities. Two of these studies permit comparisons with the present data at the level of specific measures of nonwork activities: Hanhart (zero association for club membership versus significantly positive correlations in the present study for number of organizational memberships; zero association for participation in active sports events and for attendance at sports events versus zero association in the present study for sports & games); and Larrue (zero association for reading versus zero association in the present study for intellectual activities [viz. reading]; zero association for spectator sports versus zero association in the present study for sports & games).

Cell nine, the final cell, cannot be explored in the present study. As noted earlier, the relevant literature supports the positive hypothesis for the relationship between subjective reactions to work and leisure; and among husbands, for the relationship between subjective reactions to work and family life.

Variations in Demographic Subgroups

Examination of various demographic subgroups uncovers few appreciable variations in correlational patterns. Sex differences are a case in point. For a number of cells which produce definitive correlational trends, sex makes little difference: no effect in cell one on the positive correlations between degree of involvement in work (measured subjectively) and in nonwork; no effect on the corresponding positive correlations in cell two; no overall effect on the positive correlations in cell five matching types of activities on and off the job; and no effect on the positive correlations in cell eight between subjective reactions to work and types of nonwork activities. Sex does make a difference, however, in cell seven in which the positive correlations between subjective reactions to work and degree of involvement in nonwork are larger for women than men.

In addition, the literature (though not, of course, the present data) suggests that the positive relationship in cell nine between subjective reactions to work and family life is stronger for men than women.

Sex differences aside, only one other demographic subgroup stands out with any consistency across the cells in Table 1. Nonmarried males exhibit correlations between work and nonwork that are not as positive as for other subgroups. This applies in particular to cell one, using subjective measures of degree of work involvement (weaker positive correlations); cell two, again using subjective measures of degree of work involvement (zero correlations); cell five, as regards matched activities (slightly negative correlations); and cells seven and eight (substantially negative correlations).

Several interpretations are possible for the paucity of consistent demographic differences. Taken at face value, the results indicate that the relationship between work and nonwork changes little across demographic subgroups of the population of workers. By implication, the results also tend to cast doubt on theories which predict that a particular relationship between work and nonwork holds only among a certain demographic subpopulation of workers. The pattern of demographically consistent results likewise challenges the view that the relationships observed between work and nonwork are spurious and in fact attributable to confounding by demographic variables.

Causal Inferences

As noted in Part 1, correlations between work and nonwork are compatible with two major causal directions: the effect of work on nonwork; and the effect of nonwork on work. There exists the further possibility that some other factor which correlates positively with both work and nonwork may explain the positive relationship between the two concepts. For example, the mechanism of personality type cited in connection with cell one, suggests that highly motivated, intense individuals tend to become heavily involved in both work and nonwork. Cross-sectional analyses such as those in the present study, nonetheless, do not readily resolve the issue of causal direction and the ensuing discussion of causal mechanisms is no exception.

The issue of causal inferences obviously arises only when there is an effect to explain. Data in the present study generate appreciable positive relationships between the following concepts: subjective measures of degree of involvement in work and degree of involvement in nonwork (cell one); subjective measures of degree of involvement in work and types of nonwork activities (cell two); degree of mental effort required by the job and subjective reactions to nonwork (cell four); types of work and nonwork activities (cell five); subjective reactions to work and degree of involvement in nonwork (cell seven); subjective reactions to work and types of nonwork activities (cell eight). In addition, the data establish a negative relationship between degree of physical effort required by the job and degree of involvement in nonwork (cell four). Among all these findings only those for cells one, four (effort measures), five (matched activities) and seven provide direct tests of the positive and negative hypotheses. These become the four cells of greatest causal interest.

The fact that residualized scores detract only slightly from the positive findings in cell one (based on subjective measures of degree of involvement in work) implies that the resources of time, energy, health and money make at best a marginal contribution to the positive correlations in that cell. Otherwise, as Table 1 suggests, there are four possible mechanisms to explain the positive correlations in cell one: integration, personality type, skills & abilities, and cultural pressures. Integration is certainly a plausible possibility because the positive correlations in cell five suggest a matching between, and possibly an overlap of, work and nonwork activities. The mechanisms of personality type and skills & abilities, though plausible, do not receive a satisfactory test in the present analyses. One possible strategy for selecting between them involves the use of a control on time in the work role. If the correlations between work and nonwork differ little as tenure in the labor force increases, personality type seems the more reasonable interpretation; but if the correlations increase as tenure increases, thereby suggesting a gradual and cumulative process, acquisition and development of skills & abilities appears a more likely explanation. Finally, insofar as cultural pressures are thought to apply to men in particular, the absence of a sex difference in cell one casts doubt on this fourth mechanism.

Cell four offers some support for both positive and negative hypotheses. It contains positive correlations between two measures of types of work activities (always a lot of work to do, requires a lot of mental effort) and degree of involvement in nonwork. Most of these correlations disappear, however, once the scores on nonwork are residualized, thereby indicating that major resources probably account for a substantial portion of the positive correlations.

In accordance with the negative hypothesis, cell four produces a negative relationship between extent of physical effort required on the job and degree of nonwork involvement. Since this negative relationship survives the control on resources via residualization, it gives credibility to the one negative mechanism cited in Table 1 for cell four, namely, energy as a scarce resource. These findings on physical effort may at first seem paradoxical: physical effort still correlates (negatively) with degree of involvement in nonwork, even after the control via residualization of four resources, one of which is energy. Yet the apparent paradox is easily resolved. The control on energy via residualization concerns an individual's general level of energy relative to other workers. Those in jobs which require a lot of physical effort may start out their day with a normal quota of physical energy, quite possibly more than other workers, but by the end of the work day their level of energy has sunk low. In short, there is no reason why a control on overall level of energy relative to other people should eliminate a relationship that concerns what workers do at the end of a tiring workday.

Because residualization has no substantial effect on the correlational pattern in cell five (selected matched activities), no simple appeal to basic resources will explain the findings. In line with the positive hypothesis, cell five generates positive correlations for three types of matched activities: mental effort, interpersonal involvement, and supervision. It thus offers some support for the four positive mechanisms cited in Table 1 (integration, skills and abilities, learned habits, and strong preferences) although it fails to pick among them.

As predicted by the positive hypothesis, cell seven provides a consistent array of positive correlations between subjective reactions to work and degree of involvement in nonwork, based on nonresidualized scores. Yet the process of residualization eliminates the positive trend and, in fact, replaces it with a weak negative one. Major resources, then, appear to explain the positive findings in cell seven.

Conclusions

This paper has explored two broad and potentially diffuse approaches to the relationship between work and nonwork. Underlying the positive and negative approaches investigated here are the general notions of similarity and dissimilarity, respectively, between work and nonwork. The fact that most of the correlations that test the basic positive and negative hypotheses emerge as positive accords merit to the broad approach in general and to the broad positive approach in particular. The fact that certain of the correlations linking work and nonwork prove to be negative (e.g., physical effort at work and degree of involvement in nonwork) establishes that sometimes there is value to searching for contingent (or conditional) relationships.

Whereas previous studies have typically looked at work in conjunction with leisure or family life but not in terms of both concepts, the present study includes the two areas of nonwork along with the domain of work. Yet the comparison between findings for work & leisure and work & family is not without its limitations in the present dataset. Of the study's nine measures of nonwork, only one taps family activities directly (family activities), although another (domestic activities) seems close enough to group with family activities. Moreover, insofar as family life often includes leisure activities, and leisure activities frequently involve family members, the remaining comparison of seven measures of leisure activities and two measures of family life does not avoid ambiguity. Cells two and eight permit a rough comparison between the results obtained for leisure and those obtained for family life. With significant correlations as the criterion, cell two reveals a tendency for degree of involvement in work to be more positively related to participation in family life than in leisure activities, for objective and subjective measures and for nonresidualized and residualized measures. Cell eight displays a parallel tendency for subjective reactions to work to be positively associated with participation in family life than in leisure activities, for both residualized and nonresidualized measures.

It seems clear from this paper that "the long arm of the job", to use Meissner's (1971) phrase, extends to the domain of nonwork in its different forms. It also seems clear that, except where the expenditure of physical effort is involved, work and nonwork tend to resemble rather than to contrast with each other. What is not clear is why these two domains are positively related. Only when the various causal directions and causal mechanisms have been fully explored will policy makers know how to help people keep their work in step with the rest of their lives.

References

Bacon, A. W. Leisure and the alientated worker: A critical reassessment of three radical theories of work and leisure. Journal of Leisure Research, 1975, 7, 179-190.

Bishop, D. W. Stability of the factor structure of leisure behavior: analyses of four communities. Journal of Leisure Research, 1970, 2, 160-170.

Brown, R., Brannen, P., Cousins, J., & Samphier, M. Leisure in work: The 'occupational.culture' of shipbuilding workers. In M. Smith, S. Parker, & C. Smith (Eds.), Leisure and Society in Britain. London: Allen Lane, 1973.

Clark, R. A. and Gecas, V. The employed father in America: A role competition analysis. Paper presented to the 1977 Annual Meeting of the Pacific Sociological Association, 1977.

Clark, R. A., Gecas, V., and Nye, F. I. "Career widows"? Husbands' work involvement and marital role performance. Paper presented to the 1977 Annual Meeting of the American Sociological Association, 1977.

Converse, P. E. and Robinson, J. P. 66 Basic Tables of Time-Budget Data for the United States. Ann Arbor: University of Michigan, Survey Research Center, June, 1966.

Converse, P. E. and Robinson, J. P. Social change reflected in the use of time. The Human Meaning of Social Change. New York: Russell Sage Foundation, 1972.

de Grazia, S. The uses of time. In R. Kleemeier (Ed.), Aging and leisure. New York: Oxford University Press, 1961.

de Grazia, S. Of Time, Work, and Leisure. New York: Twentieth Century Fund, 1962.

Goldthorpe, J. H., Lockwood, D., Bechhofer, F., and Platt, J. The affluent worker in the class structure. London: Cambridge University Press, 1969.

Grubb, E. A. Assembly line boredom and individual differences in recreation participation. Journal of Leisure Research, 1975, 7, 256-269.

Hagedorn, R. and Labovitz, S. Participation in community associations by occupation: A test of three theories. American Sociological Review, 1968, 33, 272-283.

Hanhart, D. Arbeiter in der Freizeit: Eine sozial-psychologische Untersuchung. Bern: Hans Huber, 1964.

Havighurst, R. J. The nature and values of meaningful free-time activity. In R. W. Kleemeier (Ed.), Aging and leisure. New York: Oxford University Press, 1961.

Howard, D. R. Multivariate relationships between leisure activities and personality. Research Quarterly, 1976, 47, 226-237.

Kaplan, M. Leisure in America. New York: Wiley, 1960.

Kelly, John R. Work and leisure: A simplified paradigm. Journal of Leisure Research, 1972, 4, 50-63.

Larrue, J. Loisirs ouvriers chez les Metallurgistes Toulousains. Paris: Mouton, 1965.

London, M., Crandall, R., and Fitzgibbons, D. The psychological structure of leisure: Activities, needs, people. Urbana: University of Illinois, 1976.

Lundahl, A. Leisure and recreation: Low income investigation. Stockholm: Allmanna Forlaget, 1971.

Matthews, A. T. V. and Abu-Laban, B. Job satisfaction and leisure-time activity in a restrictive environment. Sociology and Social Research, 1959, 43, 189-196.

McKechnie, G. E. The psychological structure of leisure: Past behavior. Journal of Leisure Research, 1974, 6, 27-45.

Meissner, M. Workers and work: Meanings and methods of research. International Institute for Labour Studies Bulletin, 1970, 7, 293-314.

Meissner, M. The long arm of the job: A study of work and leisure. Industrial Relations, 1971, 10, 239-260.

Meyersohn, R. A critical examination of commercial entertainment. In R. W. Kleemeier (Ed.), Aging and Leisure. New York: Oxford University Press, 1961, pp. 243 ff.

Meyersohn, R. The sociology of leisure in the United States: Introduction and bibliography, 1945-1965. Journal of Leisure Research, Vol. 1, Winter, 1969, 53-69.

Morris, G., Pasewark, R., and Schultz, J. Occupational level and participation in public recreation in a rural community. Journal of Leisure Research, 1972, 4, 25.

Orthner, D. K. Patterns of leisure and marital interaction. Article is a revision and extension of papers presented at the meetings of the Pacific Sociological Association, March, 1974, and the National Council on Family Relations, August, 1975.

Parker, S. R. Work and non-work in three occupations. Sociological Review, 1965, 13, 65-75.

Ritchie, J. R. B. On the derivation of leisure activity types--A perceptual mapping approach. Journal of Leisure Research, 1975, 7, 128-140.

Salaman, G. Community and occupation. London: Cambridge University Press, 1974.

Seeman, M. On the personal consequences of alienation in work. American Sociological Review, 1967, 32, 273-285.

Wilensky, H. L. Life cycle, work situation, and participation in formal associations. In R. W. Kleemeier (Ed.), Aging and leisure. New York: Oxford University Press, 1961.

Willmott, P. Family, work and leisure conflicts among male employees: Some preliminary findings. Human Relations, 1971, 24, 575-584.

Witt, P. A. Factor structure of leisure behavior for high school age youth in three communities. Journal of Leisure Research, 1971, 3, 213-220.